INTERNATIONAL HELLENIC UNIVERSITY

Algorithmic modeling for the efficient detection of epilepsy disorder on EEG signals.

Charalampia Komopoulou

SID: 223240734767

SCHOOL OF SCIENCE & TECHNOLOGY

A thesis submitted for the degree of

Master of Science (MSc) in Data Science

NOVEMBER 2018

THESSALONIKI – GREECE

Algorithmic modeling for the efficient detection of epilepsy disorder on EEG signals.

Charalampia Komopoulou

SID: 223240734767

Supervisor:	Dr. Agamemnon Baltagiannis
Supervising Committee Members:	Assoc. Prof. Name Surname
	Assist. Prof. Name Surname

SCHOOL OF SCIENCE & TECHNOLOGY

A thesis submitted for the degree of

Master of Science (MSc) in Data Science

NOVEMBER 2018

THESSALONIKI – GREECE

Eyes that are for me seas of patience.

To my father Vissarion.

Abstract

This dissertation was written as a part of the MSc in Data Science at the International Hellenic University.

Epilepsy is a disorder that a considerable number of people suffer from. Even though seizures are highly researched there is no known cure for the disorder apart from the resection of the epileptogenic zone which works in a limited number of cases. The aim of modern neuroscience is to automate the procedure of predicting if an EEG signal presents seizure or healthy activity, with the highest accuracy possible. Working in that framework, what this thesis presents, is initially the research on the existing literature by displaying the best classification schemes and trials. Visualizations, plots, and tables were displayed, using R, in order to give better perception of the data. In the next step of the study a group of algorithms and classifiers were presented and tested using extensively Weka 3.8 data mining tool. *Tree* algorithms were proven to have the best performance on predicting if the signals were ictal or healthy and the average accuracy achieved by the *trees* group was 86% while the next best group had 81% average accuracy. The best accuracy received by one algorithm was 94.7% given by the Random Committee algorithm, which used as classifiers of the ensemble *decision trees*. Finally, in the study it is proposed that the *Tree* algorithmic model fits best the data and presents remarkable predictions especially if we consider that the data had undergone only limited preprocess and no normalization at all.

I would like to thank my supervisor Dr. Agamemnon Baltagiannis for the help and trust he showed me. Also, my parents and friends that stood by my side all these months.

Charalampia Komopoulou

30 November 2018

Contents

1. Introduction

In this chapter, general information on epilepsy, seizures, electroencephalogram, statistics on the disorder and useful information about several procedures and recording techniques are displayed.

1.1 What is epilepsy

Epilepsy is a neurological disorder characterized by recurrent unprovoked seizures, with various durations. It is the most common neurological disorder after stroke and it affects approximately 1% of the world's population. The population diagnosed with the disorder increases by 5% each year and nearly 80% of the suffering population comes from developing and non-developed countries as mentioned in the article (Tao Zhang, 2018).

Seizures can last from a few seconds to several minutes and they are a result of abnormal and excessive neuronal activity in the cerebral cortex. The human body responds to unexpected and uncontrolled electrical discharges in a group of brain cells by seizing. Seizures can result in several injuries for the person that undergoes them. Usual symptoms that accompany seizures are temporary confusion, staring spell, uncontrollable jerking movements of arms and legs, and loss of consciousness or awareness. In epilepsy, seizures tend to recur with no immediate underlying cause, in contrast to the isolated ones that can be caused by poisoning, head trauma, massive body injuries, extreme dehydration, and drugs or alcohol withdrawal.

The cause of epilepsy in most cases is unknown. However, some of the known causes of seizures in patients that suffer from the disorder are, brain tumor, stroke, brain injury, genetic mutations, brain aneurysm, and birth defects. People suffering from epilepsy are treated in various ways around the world. The only fact is that until now there is no clear treatment for the disorder. Most of the times, medical prescriptions are administered to smooth or delay the symptoms of seizures and help people have a better quality of life.

1.2 Categorization of epilepsy and studies

In recent years, a lot of research and studies focus on finding the exact reasons that provoke the initiation of seizures in patients that suffer from the disorder. Many seizures were categorized according to their symptoms, their expression and their localization on the triggering spot of the brain. The most common type of seizures are convulsive. Hence, there are several non-convulsive seizures that occur, but they are not widely known because they do not have such extreme symptoms on their occurrence. Some of the convulsive seizures are called generalized (NYU Langone Medical Center, n.d.) and they affect both hemispheres of the brain. Some others are called focal/partial seizures and affect only one hemisphere, but sometimes progress and become generalized. There are also the unilateral and the unclassified seizures.

According to neurologists, epilepsy is marked and characterized by two or more intermittent seizures. As mentioned in (Manish N. Tibdewala, 2017) in the generalized seizures there is loss of consciousness while in partial there maybe or maybe not loss of consciousness. The generalized seizures are also discriminated from partial because the latter is more common in very young or elderly people.

As mentioned above there are a few categories describing the spatial range of seizures. In this study though, we consider seizures as events that need to be monitored and treated but we do not care about the category of the seizure or the spot of the brain that the seizure is located or initiates.

1.3 What is the EEG and how does it work with seizure detection.

The most common and highly used way to confirm and detect epileptic seizures on patients nowadays is the electroencephalogram aka (EEG). The EEG is a record of electric stimulations generated by brain nerve cells. There are two different types of EEGs depending on the part of the brain that the signal is taken from. There are *scalp* and *intracranial* EEGs. The procedure of signal recording is different in the two cases since in the latter the signal is acquired during surgery.

In the scalp EEG the electrodes which are small metal disks created to perceive the electric signals initiating in the brain, are placed in the scalp. The electrodes are connected to an EEG monitor which creates records of the brain signals. In the intracranial EEG, the recording of the signal is taken during an operation in which the surgeon places the electrodes directly to brain tissue. Most of the times the reason that a surgical EEG is performed is to find which exact place of the brain the seizures start in or when scalp EEG gives inconclusive results.

The electrodes are of low impedance and manufactured to perceive current of small voltage. As mentioned in (Hojjat Adeli, 2003) the sensed changes in the voltage difference between electrodes are amplified, using a differential amplifier technology which takes two electrical inputs and displays the output as the difference of the two inputs. Then the signal is transmitted to an appropriate for this purpose computer program which displays the tracing of the voltage potential recordings. In that way, a continuous graphic exhibition of the spatial distribution of the changing voltage field over time is displayed on the EEG. (Hojjat Adeli, 2003)

The most common method for placing the electrodes in the scalp is the 10-20 system. In this system, the scalp is “separated” in different regions based on a 10%-20% distance between electrodes. Since the implementation of the electrodes has a determinant role on the detection of seizures, medical society defined the 10-20 system as the state-of-the-art system for electrodes placement during an EEG.

Neuroscientists defined two main factors for false detections or no detection on the scalp electroencephalogram with the area of the cortex involved to the seizure being the first one. It was observed that when the area of the seizing part of the brain is small, it is highly unlikely that the EEG will catch abnormal stimulation. In average the seizing area should be between 10 and 20 cm^2 so that the EEG perceives abnormalities. The second factor is the localization of the seizure. The EEG electrodes perceive seizures and abnormal stimulations when the area of the seizure is near to the scalp. For that reason, it is proven that it is very difficult to detect any signal when the area of the seizure is located deep in the brain. In those cases, the scalp EEGs are inconclusive, and an intracranial EEG is preferred.

Epilepsy is detected using the electroencephalogram, however, a simple recording is not enough to confirm the existence of seizures as a disorder in a patient. Further biochemical testing needs to be done to confirm the epileptic seizures as a permanent condition. Clinically in order to diagnose and predict epileptic seizures in patients, brain activity needs to be monitored through EEG signals which contain patterns and markers that declare the existence of abnormal brain activity.

In the EEG of a patient two abnormal states are detected, the ictal state in which the patient is seizing and the interictal state in which the patient experiences some symptoms of seizure that are not convulsive. The interictal state is usually detected before the initiation of a seizure and many times after the experience of a seizure. The usual duration of the interictal state is from a few seconds to several minutes. Hence, many epileptic patients experience long interictal states which last more than 45 minutes. During that time the patient might feel fuzziness, disorientation, loss of speech, blurry sight and extended loss of senses (Mayo Clinic , n.d.). As mentioned in (Rajendra Acharya. U, 2015) the ictal signals are continuous waves which exhibit spikes and sharp wave complexes. In the interictal state, the received EEG signal is showing transient waveforms and sharp and spiky waves.

Most of the times the EEG records interictal states as it is difficult to obtain a seizure at the specific time that the patient gets the electroencephalogram. Clinicians rely on the signals of the interictal states to diagnose epileptic patients as they rarely acquire EEGs that display seizures. For the above reason, the duration of the EEG recordings needs to be long so that clinicians have enough monitored signal to analyze and localize the normal, interictal and predict or rarely visualize ictal episodes.

The epilepsy condition is detected by trained neurophysiologists who study and visually inspect long duration electroencephalogram signals. As mentioned in (Hojjat Adeli, 2003) the signals require visual inspection in which the clinician examines features like voltage, amplitude, frequency, waveform regularity, and measures the reactivity to hyperventilation, photic stimulation and eye-opening by specific markers that appear in the signals. The neurophysiologists also need to identify temporal abnormalities and spatial range of the signals. The procedure though is time consuming since, until the recent years the clinician needed to manually study and find the abnormal patterns in continuous and long duration signals. The aforementioned way to detect epilepsy, apart from time consuming was also susceptible to the clinician’s inspective skills, concentration, and subjective judgment. In order to overcome the above obstacles and limitations, the need to automate the procedure of the visual inspection and detection of abnormalities is arising. For that reason, there has been a lot of research and studies on the automation of the processes.

Most of the already existing algorithms and automation techniques are data dependent and thus present high overfitting rates. The purpose of modern neuroscience on the field of seizure identification is to create a universal method to process EEG signals and discriminate recordings regardless of the dataset. The main problem though is that the seizure signals of patients present unique spectral and spatial characteristics that cannot be used to identify seizures in other patients. However, there have been groups of characteristics and recurring patterns that appear in most cases and they are currently used for the identification of seizures. For the above reason, a lot of research and algorithmic modeling trials have been conducted on the specific topic and even if there is not a generic approach yet, several solutions and models are proven to have good applicability and efficacy in a wide range of datasets and EEG signals.

1.4 EEG signal collection process

The electroencephalogram is a procedure not usually performed. There must be some indication of seizure activity or brain abnormality, or the patient may have already had a seizure, in order for a physician to order an EEG. For that reason, there is not enough information not only on seizure signal datasets but also on healthy subjects since healthy people do not have EEGs. That is the main problem in the effort to automate seizure detection processes. Usually, the EEGs are recorded in hospitals after a patient has experienced convulsive events or after the diagnosis of the disorder. EEGs are recorded every 6 to 12 months in patients unless there the physician recommends that there is a reason to perform an EEG sooner.

During the EEG the patient may be conscious or sedated depending on the kind of EEG. There are also several consciousness levels like full cognitive state with eyes open or eyes closed, under the influence of ethanol state in which the patient is semi-sedated and finally fully sedated state. All the above cases present different signals in terms of both spatial and spectral range since subjects have a completely different stimulus.

The EEG monitor can record brain activity for a few hours to a few days. In most cases, a few hours of recording are enough for the neuro-physician to study and form an opinion on the state of the patient. The purpose of a long-duration EEG is most of the times to catch ictal and inter-ictal stages. Though many times the EEG recordings are not useful since there are cases of seizures that are not spotted by the EEG monitor. As mentioned in (U. Rajendra Acharya, 2017) patients that their seizures emanate from a very small area or a very deep spot in the brain or initiate in the frontal lobe, may present normal EEG signals while undergoing a seizure. In the above cases, the intracranial EEG is preferred because it has higher possibilities of catching the seizures but it’s an invasive approach and even that procedure does not guarantee that the monitor will record the abnormalities of seizure signal.

2. Literature review

In the literature review chapter, the collected and meticulously studied scientific material on classification algorithms and processing techniques of EEG signals are presented along with information on the methodology used in this part of the thesis and a short summary.

2.1 Methodology

For the sake of this study, a lot of research has been conducted on fields relevant to medicine, EEG signals analysis, seizure detection algorithms, and time series analysis. In order to gather useful information and get an image of the already existing work and studies, a few online libraries and databases were used. In the first step of the research it was obvious that a huge amount of content was relevant to this thesis field of study, and for that reason, some inclusion criteria were implemented. Articles, conference papers, and some dissertation thesis, all published after 2000 were included in the literature review.

The research process started by searching in Science Direct, IEEE and ELSEVIER online libraries. The keywords applied in order to retrieve relevant articles were: signal, seizure detection, algorithms, EEG and it should be mentioned that these words were combined in phrases using extensively the logical AND, in an effort, to increase the number of relevant articles appearing in the search. The phrases used in the search of the libraries were the following:

→ "EEG signal" AND "seizure detection" AND "algorithms"

→ "seizure detection" AND "electroencephalogram signal" AND "machine learning"

→ "time series" AND "signal analysis" AND "seizure prediction"

The number of relevant to the above keyword phrases articles was initially 1500. After reading the titles of those articles, checking for duplicates and eliminating the ones published before 2000, the number of articles reduced to 200. In the next step, the abstract of the articles was studied, and a short and quick reading of the articles was performed leading to 105 articles to read in full-text mode. In the final step, after studying the remaining articles in full-text mode 34 of them were used in the literature review. The described process can be seen also in the following diagram in figure Fig. 1.

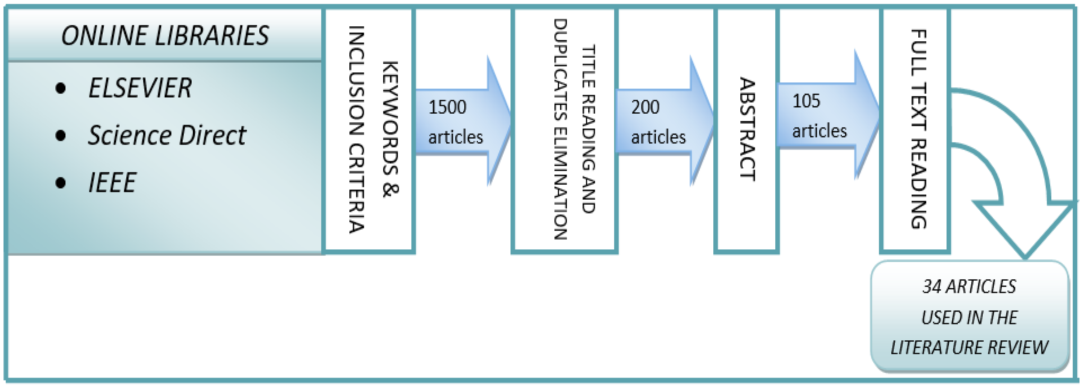

Fig.1

2.2 Feature extraction techniques.

As mentioned before the duration of the EEG signals is usually long so that clinicians and computer programs have enough information in order to detect and decide if there is abnormal brain activity. Thereby, the necessity to find recurring features and marks that characterize the non-ictal and ictal brain signal is increased. For the above reason several different features were found and used in the examined literature in the effort to store and encode useful information about the different patterns and markers of ictal and healthy brain activity.

The empirical mode decomposition and the discrete wavelet transform were the most common techniques used for the extraction of significant and efficient features. However, there were several other techniques (17 in total) as it can be seen in the following figure. (Fig.2)

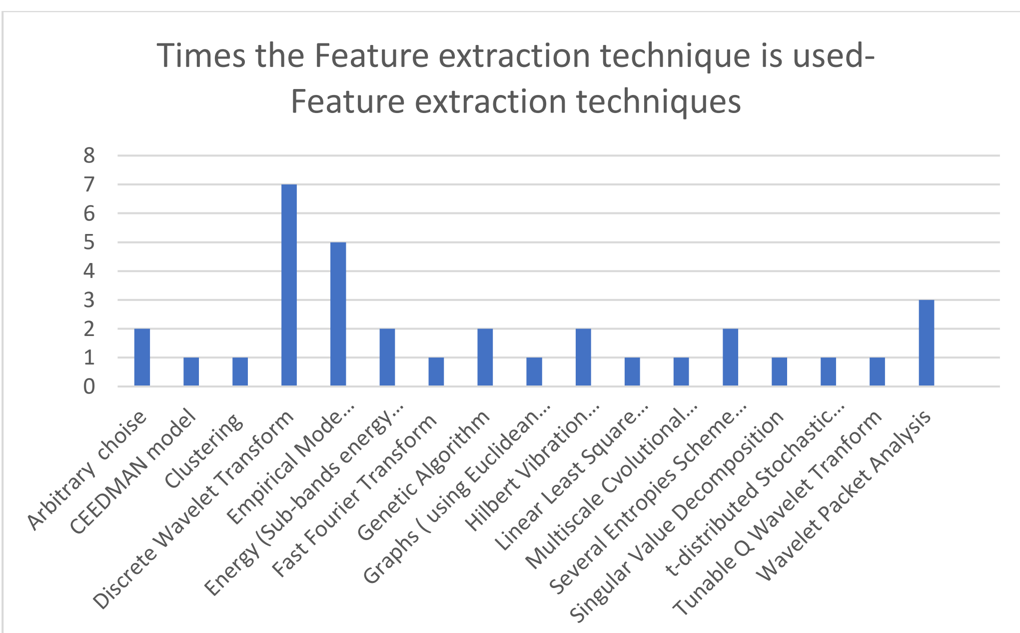

Fig.2

The horizontal axis of Fig. 2 displays the different feature extraction techniques found in the literature and the vertical axis shows how many times the specific extraction method was found in total.

It should also be mentioned that in many articles more than one feature extraction techniques were implemented. Furthermore, almost in all the examined papers and studies, it was noted that the feature extraction process was proven to be extremely crucial for the detection of seizures and the performance of the classifiers used for the final predictions. For those reasons, a lot of emphases is given by many authors in the criteria used for the elimination of features and selection of the most contributing ones in the trials and model tests.

2.3 Terminology

For the efficient study and understanding of this paper, in the section below, some of the most frequently used terms are defined and explained.

The epileptogenic zone was defined by (Lüders HO, 1993) as "the area of cortex that is necessary and sufficient for initiating seizures and whose removal (or disconnection) is necessary for the complete abolition of seizures".

Temporal lobe was defined by (Miller, 2003) as "the part of each side or hemisphere of the brain that is on the side of the head, nearest to the ears."

Hippocampal formation defined in (Mosby, 2009) is "an imprecisely defined structure in the medial temporal lobe of the brain, which is thought to play a central role in memory, spatial coding and in control of attention."

Wavelet Analysis Is a technique used for feature extraction and according to (Hojjat Adeli, 2003) there are two types. The Discrete Wavelet Transform and the Continuous Wavelet Transform. Discrete Wavelet Transform which is used in most of the studies, is a weighted sum of a series of base functions. Specifically, the inner product of the EEG signals is recorded in different points and the results are presented as the sum of the series. The DWT decomposes a signal into approximation and detail coefficients at the first level. Afterward, the approximation coefficients are decomposed into next level of approximation and detail coefficients. In that way, the DWT can reveal details on the examined signal in both time and frequency domain accurately.

It should be mentioned that the most significant element of wavelet analysis is the wavelet function which is a continuous, zero mean amplitude and finite function. Wavelet analysis is highly recommended as a scheme useful for feature extraction when the signals are non-stationary. Thereby it is considered as an effective way for feature extraction in the case of EEG signals which are known to be non-stationary, continuous and time-dependent.

Entropy according to (Rajendra Acharya. U, 2015) is a measure of uncertainty. High entropy rates mean that there is high unpredictability on the data and low entropy corresponds to high regularity. It is a non-linear index showing the degree of chaos within a system and is highly used in signal analysis. The entropies are categorized into spectral and embedding groups. Entropies that depend on the amplitude value of the power spectrum of a signal are called spectral and the ones calculated using time series are defined as embedding. (Rajendra Acharya. U, 2015)

Approximate Entropy is the logarithmic likelihood that the trend the data patterns are close to each other will remain close in the next comparison with another pattern. It is also a measure of data regularity. (Rajendra Acharya.U, 2012)

Sample Entropy, proposed by Richman and Moorman is the negative natural logarithm of an estimate of the conditional probability that patterns of length e that match point-wise within a tolerance r also match at the next point (Sample Entropy, n.d.). It is also a measure of regularity and is highly dependent on the length of the processed signal.

Linear Discriminant Analysis assumes normal distribution of the data, with equal covariance matrix for the two classes. The technique has low computational requirements which renders it suitable for online and real-time classification problems such as EEG signal analysis. (Garrett et al., 2003). (Ling Guo, 2010)

Support Vector Machine which is one of the classifiers spotted in several studies is classifier uses a discriminant hyperplane to identify classes. Each time the selected hyperplane is the one that maximizes the margins, for example, the distance from the nearest training points. As (Blankertz et al., 2002) mention in their study, "maximizing the margins is known to increase the generalization capabilities ". (Ling Guo, 2010)

The Bayesian classifier aims at assigning to a feature vector the class based on the highest probability. The Bayes rule is used to compute the a posteriori probability that a feature vector has when belonging to a class. Using the MAP (maximum a posteriori) rule and these probabilities, the class of this feature vector can be estimated (Fukunaga, 1990). (Ling Guo, 2010)

Nearest neighbor classifiers are relatively simple. They assign a feature vector to a class according to its nearest neighbor(s) and they are discriminative non-linear classifiers (Garrett et al., 2003), (Ling Guo, 2010)

Empirical mode decomposition (EMD) is a data-dependent approach quite common and suitable for decomposition of non-linear and non-stationary signals into symmetric, amplitude and frequency modulated (AM–FM) components known as intrinsic mode functions (IMFs). (Rajeev Sharma, 2015)

Cross-correlogram is a tool used to measure the randomness in a dataset. This randomness is verified by computing autocorrelations for data values at different time lags. If the data is random the autocorrelations will be near zero, otherwise the autocorrelation values will vary. (Friendly, 2002)

2.4 Algorithmic methods and studies results

The study of literature concerning epilepsy as a disorder was proven to have a vast amount of available information. After the selection process explained in part 2.1 Methodology, the number of papers used in the review were 34 and most of them presented similarities in the processes of prediction and classification. In the following part of the literature review, some of the most relevant and useful studies and results are presented accompanied by information on the classification methods and the preprocessing steps that were followed in each experiment.

2.4.1 Studies Review

In a relatively recent study, (Javad Birjandtalab, 2017) presented the advantages of channel selection and nonlinear dimension reduction for accurate automatic seizure detection. The authors extracted some frequency domain features from full-channel EEG signals. Afterwards, they used a random forest algorithm to find which channels have higher contribution in discriminating ictal from non-ictal events. A non-linear dimension reduction technique was applied in order to find a relationship between data elements and map them in a lower dimension. For the classification process, K-Nearest Neighbor (KNN) classifier was implemented in order to discriminate seizure from non-seizure signals. The results of the experiments have shown that for 23 patients the predictions of the algorithms and their accuracy were better than other visual and manual detection techniques. In that process, power spectral analysis was used to extract features in each time window and for every channel and every subject. The best and most contributing channels were used in the classification. That method provided robustness to noise and used the most useful channels to classify the wanted segments. As noted in the article (Javad Birjandtalab, 2017), they used a “t-distributed stochastic neighbor embedding (t-SNE) to embed and represent data in a lower dimension by preserving the relationship of data points in high dimensional feature space”. At last a KNN classifier was implemented in order to discriminate seizure from non-seizure events.

Apart from the full-channel EEG signals, in (Javad Birjandtalab, 2017) some limited-channel electroencephalogram recordings were used and presented several advantages. In a computational manner, the limited channel EEG simplified the complexity of detection of seizures and led to faster run time and lower computational costs, thereby making seizure detection models faster and more efficient. It also increased the accuracy of prediction in some cases by eliminating redundant and non-useful channels in the classification. In the examined case in which the channels from 23 were reduced to 1-3, made the EEG monitoring much easier and practical. In order to choose the best features to use for the classification, the Random Forest algorithm was implemented which is known to provide very good results when it comes to discarding large number of irrelevant features.

The nonlinear dimension reduction method used in (Javad Birjandtalab, 2017) was based on the t-distributed Stochastic Neighbor Embedding (t-SNE) with purpose to represent the data points into lower dimensions. Finally, the features of the 3 most important channels were represented into two dimensions. For the classification part, the authors separated several hours of EEG signals for each patient into 10 seconds segments in order to catch even the shortest in duration seizures, which last on average 10 seconds. After the channel selection and dimension reduction, each segment window was represented by only two features. The authors used then a K-NN classifier to discriminate ictal and non-ictal stages using each time the specific patient's data. And that was done because the expression of seizure signal in different patients is also different as mentioned in neurophysiology. So, the classifier was trained in a personalized manner and that increased the accuracy results on the final predictions. Since not all channels were relevant to the detection of seizures, it was for sure that some brain activity was missed in the process. Though it was certain that the missed part had no contribution in the classification part as it mostly contained information irrelevant or redundant. Finally, it was observed that the t-SNE successfully preserved the local neighborhood information and at the same time preserved the accuracy during the dimension reduction process.

In another study of nonlinear time series analysis of electroencephalogram signals for epilepsy diagnosis, the authors of (Rajendra Acharya.U, 2012) used Entropy as a measure of randomness in EEG signals. This approach had already been successfully used for time series analysis in the past. The authors applied four entropy measures to discriminate ictal, pre-ictal and non-ictal stages in EEG signals which were then "fed" into seven different classifiers. The entropies used where Approximate Entropy (ApEn), Sample Entropy (SampEn), Phase Entropy (S1) and Phase Entropy 2 (S2). The used classifiers where Support Vector Machine (SVM), Probabilistic Neural Network (PNN), Gaussian Mixture Model (GMM), Fuzzy Sugeno Classifier (FSC), K-Nearest Neighbour (KNN), Naive Bayes Classifier (NBC) and Decision Tree (DT). The performance of the used classifiers was assessed by the measures of specificity, sensitivity, accuracy, and p-value. By using

Analysis of Variance test a significant p-value was obtained (p-value<0.0001) which indicated that the features were useful for the discriminating process. Spectral entropy had a wide range of values, and it was observed that low values correspond to non-ictal or pre-ictal signals while higher values correspond to ictal signals. This is because seizure signals have a wider range of dominant frequencies. The ApEn and SampEn entropies gave a sense of the regularity of the time series signal. Both measures had higher values when the examined segment was non-ictal and lower on the other two cases (healthy and inter-ictal). The Phase Entropy 1 (S1) and Phase Entropy 2 (S2) measures presented a range of values depending on the probability density function that the bi-spectrum follows. Fuzzy Sugeno Classifier (FSC) was proven useful in this case because it took imprecise observations like the EEG signals as inputs and created precise values for outputs. It was also observed that the FSC created separable clusters for the different states of the signals and performed better in terms of classification accuracy on signal recognition. The highest accuracy received from the classifiers was 98.1% given by the FSC as said before, and the procedure was performed using 3-fold cross-validation. The non-linear entropy features used on this paper gave better results in terms of discrimination between the classes and finally, it should be noted that the proposed method achieved good automation and diagnostic accuracy.

Widely used methods for feature extraction and classification were implemented in the article (Ocak, Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy, 2009). The Approximate Entropy and the Discrete Wavelet Transform (DWT) were used for the analysis of the EEG signals which was accomplished in two steps. Firstly, the signals were decomposed into detail and approximation coefficients using the Discrete Wavelet Transform and in the next step, the Approximate Entropy for the coefficients was computed. The ApEn of seizure and non-seizure signal was totally different, increasing in that way the discrimination accuracy and leading to classification accuracy of 96%. The dataset used, contained 10 patients, half of them suffering from seizures and half healthy. The preprocessing part of the experiment in which DWT was implemented played a significant role in the accuracy score since without the DWT preprocess was found to be 73%. During seizures, the author observed that the Ap. Entropy had lower values compared to the non-seizure rates. This was an indication that non-ictal signals are not as clear to spot as the ictal signals. The data was further analyzed and tested for non-linearities in the signals, and it was found that seizure signal presented a lot of non-linear patterns. Also, the normal signals seemed to follow a Gaussian linear stochastic process. Finally, the proposed decomposition and analysis gave very good accuracy results in the specific problem.

A different article proposed the artificial neural network models for the classification of multichannel EEG signals. The whole process included three steps. At first, six features were fed in two different perceptron classifiers in order to create groupings based on the peaks appearing on the EEG signals. The groups contained epileptiform transients, non-epileptiform transients and possible epileptiform and possible non-epileptiform transients. (Nurettin Acır, 2005) As mentioned in (Tedrus GM, 2012) epileptiform transients such as spikes and sharp waves are the interictal markers of a patient with epilepsy and are the EEG signature of a seizure focus. Usually, an epileptiform transient has peak and pointed duration of 20 to 70 microseconds.

The pre-classification performed in the step described above reduced the computational time and increased the performance of the whole classification experiment. In a second stage, a nonlinear neural network was used to discriminate the cases that were not clearly classified in the first stage. Different networks were used as post-classifiers and their performance was compared to each other and evaluated. Finally, the multichannel information was all integrated for the final identification of epileptiform events on the EEGs. The visual inspection of 19 channel EEG records from 10 patients suffering from epilepsy indicated that the best performance (sensitivity 89.1%) was obtained when a radial basis support vector machine was used.

The features used to the pre-classifier in (Nurettin Acır, 2005) were: first half wave amplitude, second halfwave amplitude, first half-wave duration, second half wave duration, first half wave slope, and second half wave slope. In addition, the evaluation process contained the false detection per hour which is a measure that shows the performance of the system by giving the number of falsely determined segments of signal. The statistical measures used to evaluate the pre-classification stage were selectivity, sensitivity, and false rate. The pre-classifier gave sensitivity rate of 100%, selectivity's lowest rate 9.6% and the hourly false rate of 3371.9. This meant that the pre-classifier found only the definitive cases of epileptiform and non-epileptiform transient cases. In the third stage seizure events were identified through the integrated multichannel information. The main classifier used was the RB-SVM which gave the best results in terms of specificity, sensitivity and false rate. Selectivity now was found to be 85% which is a huge difference compared to 9.6% in the initial classification stage. The sensitivity rate was a bit reduced (89.1%) and the false predictions rate became 7.5 per hour. Several other classifiers were used in that stage like Radial Basis Function Network (RBFN), Autoregressive RBFN, the Multilayer Perceptron which all gave similar results.

In the mentioned study, 29 patients contributed and 19 of them were included in the training part and 10 in the testing part. The recording of EEGs was done while the patients were at restful wakefulness stage, and two clinicians studied the EEG signals of the epileptiform transient patients, and when their diagnosis did not meet the EEG signal was used as background signal in the process.

In another useful article, high classification accuracy was achieved through the usage of highly discriminative features which captured the nonlinearity of the EEG signals. The SVM classifier implemented by the authors of (Lina Wang, 2017) gave the very good results in terms of accuracy and rendered the classification of the EEG signals a successfully automated procedure. In the paper a wavelet threshold denoising method was implemented, accompanied by feature extraction, principal components analysis and non-linear analysis for dimensionality reduction. After experimentation, the authors observed the classification accuracy of 99.25% which was obtained after the usage of several features from various domains like time, frequency, time-frequency, and nonlinear features, fact that outperformed the single domain feature extraction methods.

Apart from the SVM which is highly used and efficient classifier when it comes to EEG signal discrimination, the Artificial Neural Network (ANN) is considered to be an equally useful forecasting algorithm. As mentioned by the authors in (Ling Guo, 2010) ANN consists of a parallel highly interconnected structure of non-linear processing elements and is highly adaptive to given inputs. The Multilayer Perceptron Neural Network is usually used in the field of medicine and especially on signal processing and pattern recognition. In this work of (Ling Guo, 2010) a three-step method was implemented. Initially, the signals were submitted into Discrete Wavelet transform process and several sub-bands of different frequency signals were created. Afterwards, the line-length feature of each sub-band was extracted and finally, the extracted features were fed into the artificial neural network to be discriminated into the two classes of ictal and non-ictal. The dataset consisted of 500 EEG segments and which were allocated to tree classes. The accuracies obtained from the Multilayer Perceptron indicated that the classifier was the best choice when compared to accuracies of other classifiers. Specifically, in the two-class classification problem, the ANN gave accuracy of 100% accompanied by very good rates of specificity (100%) and sensitivity (99.40%).

In another study, the authors of (Ahnaf Rashik Hassan, 2016) used a new signal decomposition scheme called Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEDMAN). From that scheme, six spectral moments were extracted, and the corresponding features were implemented into an ensemble learning classifier. The used classifier is called Linear Programming Boosting (LPBoost). Analysis of the features and classification algorithms have shown that the specific scheme presented superiority against others in terms of accuracy, Cohen's Kappa coefficient, sensitivity, and specificity. Several spectral features were used to capture the discriminating factors of each segment of EEG signal. The moment based spectral features extracted from the CEEDMAN model and used in the study were spectral decrease (SD), spectral roll-off (SR), spectral spread (SS), spectral flatness (SF), spectral centroid (SC) and spectral slope (SSl). In addition, the features were extracted from different models and were submitted into hypothesis testing which reduced the computational burden and ensured the success of the features in the classification process. Some of the models used to "generate" the spectral features were: Empirical Mode Decomposition

(EMD), Ensemble Empirical Mode Decomposition (EEMD), CEEDMAN, Discrete Wavelet Transform and the Dual-tree Complex Wavelet Transform (DT-DCT). Though the performance of CEEDMAN is the best. Except for the several features, in this paper, two signal analysis schemes for the inspection of the performance of the implemented models were used. The first one was discrete wavelet transform (DWT) and the other one was dual-tree complex wavelet transform (DT-CWT). Fed with the CEEDMAN features the LPBoost algorithm produced very good classification results and ensured high accuracy and therefore rendered the automatic seizure detection process robust and easily usable. Additionally, the LPBoost algorithm achieved convergence using a small number of learners which were not allocated in advance, and thereby reduced the computational cost and the total run time. The proposed method was single channel based and suitable for real-time classification of EEG signals.

A study by (Ozan Kocadagli, 2017) presented several classification schemes and classification models with some of them being Artificial Neural Networks (ANN), Gradient Based (GB) algorithms and Genetic Algorithms (GA). For the extraction of useful features, the discrete wavelet transform was used and for the dimensionality reduction of the feature space the fuzzy relations model was implemented. The ANN classifier was trained using the gradient based algorithm along with several others. Cross validation and information criteria were implemented in the process, ensuring quick convergence of the classifier. Apart from the other processes, in an effort to choose the best, in terms of efficiency, number of neurons in the hidden layer of the artificial neural network, some information criteria were imposed by the authors. The ANN's cross entropy and mean square error brought out more reliable and robust models in terms of classification, accuracy and model complexity. Also, to ensure better classification, the EEG signals were partitioned in segments of different bandwidth, and approximation and detail coefficients were also extracted through discrete wavelet transform process. Useful features that express the non-linear and dynamical structures of the signals were fed into the classifiers and judging from the results, ensured the efficacy and accuracy of the classifier. Though, since the dimension of the feature space in the case of EEG signal processing was big, dimension reduction was performed using fuzzy relations. In that way, only the significant features were kept, and the complexity of the whole process was reduced. Finally, to detect epilepsy the selected signal segments were processed by the artificial neural networks-based information criteria and cross-entropy. In that way, deep signal analysis was managed and best classification results were achieved. In order to eliminate overfitting, the authors decided to split the data into three sets, namely: train, validation and test set. Finally, the algorithm was programmed to stop when the error rate on the validation dataset increased.

Another study which deployed a genetic algorithm and a wavelet transform scheme for classification of EEG signals was implemented by the author of (Ocak, Optimal classification of epileptic seizures in EEG using wavelet analysis and genetic algorithm, 2008). The genetic algorithms have been previously implemented on (Ozan Kocadagli, 2017) with average results when compared to other techniques implemented on the same paper. Concerning the process of the classification, initially, the signals were decomposed into sub-bands of different frequency through a 4^{th} level wavelet packet decomposition. The Approximate Entropy of the wavelet coefficients was calculated afterwards and used as a feature in the classification process. The genetic algorithm found the most important features for the classification and discriminated the ones that gave important accuracy lift in the predictions. Finally, the important features were used by the Learning Vector Quantization classifier which produced the prediction for the class of the signals. After trials on the scheme it was shown that the best classification accuracy achieved, was on average 96% which was considered a good rate, but highly correlated and dependent on the selected features, as the authors themselves mention.

Another interesting three-class classification problem was presented by (Mingyang Li, 2017) in an article using many classification models in order to accurately allocate the dataset's EEG signals into ictal, non-ictal and interictal. Some of the used by the authors schemes were wavelet-based envelope analysis (EA) and neural network ensemble (NNE). Neural Network Ensemble (NNE) was considered a method highly researched in the field of signal processing and seizure detection. It establishes good precision and generalization rate, ensures computation simplicity since the used neural networks are independent of each other and achieves efficient analysis in complex and long duration brain signals by providing high accuracy results. In order to extract significant features, the discrete wavelet transform was implemented along with the envelope analysis method. The nature of signals recorded from EEGs was considered as complex, random, non-linear, and non-stationary, and thereby there was a need for preprocessing of the signals, significant features extraction, and classification performance testing. As said before discrete wavelet transform was used to decompose the signals into discrete sub-bands and envelope spectrum of the sub-bands was acquired using the Hilbert Transform. The feature extraction process included the mean, energy, standard deviation, and maximum value of the envelope spectrum in each subband, and also the mean, energy, standard deviation and maximum value of the raw electroencephalogram signals. Envelope analysis is an efficient method for detection of periodic impact of the signals. In this paper, the envelope spectrum was created using the Hilbert Transform which created smooth and clear curves and thereby provided excellent facility and practicability. The most significant features that were extracted from the envelope spectrum were then fed into the neural network ensemble classifier.

In the paper of (Mingyang Li, 2017) a new neural network ensemble model designed for epilepsy detection was implemented. The proposed scheme achieved considerably better accuracy results 98.78% comparing to other methods**.** In the feature extraction process, several methods have been applied but the most common and efficient was the wavelet transform which provided simultaneous frequency and time views of signal segments and achieved to capture transient features like those of the epileptic seizures on EEG signals. Envelope analysis which has already been used in neonatal signal analysis was tried in general signal processing but in this paper, the authors persisted on wavelet transform with envelope analysis to extract the significant features. The Hilbert Transform was implemented by (Mingyang Li, 2017) in order to demodulate the sub-bands created by the wavelet transform and then envelope spectrum was calculated for each sub-band. The proposed Discrete Wavelet Transform method for feature extraction achieved a higher classification accuracy compared with Empirical Mode Decomposition-based feature extraction. The Neural Network Ensemble showed superiority in dealing with three-class classification problems and the neural net are completely parallelized and uncorrelated to each other. The DWT method based on envelope analysis outperformed every other model in terms of precision and showed that the used features increased the discrimination performance of the whole scheme. The highest accuracy achieved from the NNE model and with the DWT using envelope analysis for feature extraction was 98.78% which was a considerably good classification rate. In order to compare and have an opinion on the performance of the specific model, several other classifiers like KNN, SVM, and others were tested and also some other feature extraction methods were implemented (DWT, DWT with HT, Empirical Mode Decomposition and others). The next best accuracy spotted during the experimentations was acquired by the K-Nearest Neighbor classifier and it was 95.33%. Two statistical measures were used to assess the superiority of the used networks and that were sensitivity and specificity. Finally, the proposed scheme had many advantages and among them, the most important was the utilization of discrete wavelet transform fusion with the envelope analysis as feature extraction method and the employment of the ensemble neural networks. The algorithmic model provided excellent results and was considered to be suitable for extensive clinical validation because of its superior performance and stable structure.

The Empirical Mode Decomposition which is a method extensively used in many papers dissolves the EEG signals into intrinsic mode functions (IMF) that are oscillatory and symmetric. The IMFs extracted from the decomposition of the signals are displayed in decreasing order of frequency, in which the first component is linked with the highest frequency. In (Rajeev Sharma, 2015) the authors used Intrinsic Mode Functions for the extraction of features since EEG signals are non-stationary and non-linear. It was shown that the Intrinsic Mode Function (IMF) created a symmetric feature space which was useful for the efficient classification of EEG signals. In the study, new features were introduced on 2-Dimension and 3-Dimension Phase Space Representation (PSR) of signals with purpose the classification of ictal and non-ictal signals. Some prefixed values of time delay and embedding dimension were suggested for the creation of the phase space representation (PSR) of the intrinsic mode functions. The input features used in the least square support vector machine were a 2-D phase space representation under a 95% confidence interval and the interquartile range of the Euclidean distances of the 3-D PSR vectors. The authors decided to use three different kernel functions and performed several experiments using the Least Square-SVM. Namely, the kernel functions were: Mexican hat wavelet kernel, Radial Basis Function, and the Morlet wavelet kernel and all of them were evaluated together with the LS-SVM classifier. It should be noted that kernel function and features were chosen using trial and error method. It was observed that with 4000 window samples and the Morlet wavelet kernel classification accuracy rate was 98.67%. When the window size of the signals was reduced to 2000 sample window, the accuracy dropped slightly at 97% and finally when the window size was 1000 or 500 samples, the accuracies dropped more at 96.33% and 95.67% respectively. So, it was observed that the bigger the window sample, the better the rate of accuracy acquired by the LS-SVM classifier. The Morlet kernel function and the LS-SVM provided the best recorded measurements for paper (Rajeev Sharma, 2015), in terms of accuracy, sensitivity, specificity, negative predicted value, positive predicted value, and Matthews correlation coefficients, which all had rates between 96% -100% for a window size of 4000 samples.

A classification experiment different from the others was presented by the authors of (Polat, 2007) which applied a hybrid system based on Fast Fourier Transform and the Decision Tree Classifier. The initial purpose of the study was to identify the features that would be then used by the classifier using for that purpose the Fast Fourier Transform (FFT). After that, the decision tree classifier was implemented in order to make decisions on the class of the EEG signals. Ten-fold cross validation was used in the dataset and several measures like sensitivity, specificity, and accuracy were extracted. The best accuracy acquired from the decision tree classifier was 98.68%. A quite uncommon thing spotted in (Polat, 2007) was that a huge number of 129 features was obtained from the Fast Fourier Transform implementation and all of them were applied to the classifier. The feature extraction process was carried out, using Welch's method which is considered a classical non-parametric method of spectrum estimation based on the Fast Fourier Transform. Finally, the authors spotted that there were several advantages of this hybrid implementation system among which were: rapid classification results, robust allocation of the signals in the corresponding class, easy implementation of the whole process and at last, non-invasive and cost-effective way to automate seizure detection process.

Another novel method for the detection of epileptic seizures on EEG signals was proposed by (Mohammed Diykh, 2017). At first, the EEG signals were imposed into dimensionality reduction in order to eliminate irrelevant, non-useful information and also reduce the computational time of the whole process. As mentioned in other papers, EEG signals are periodic, non-linear, non-stationary and contain redundant patterns. Dimensionality reduction was found to be one of the factors that increase the accuracy of the classifiers and reduces the complexity of the executed models. In that framework, (Mohammed Diykh, 2017) split the signal segments into 32 clusters chosen dynamically in each training session and extracted statistical measures in each stage of segmentation. The segmentation process was programmed to stop when there was no improvement in classification accuracy. After some processing, it was decided that 12 features would best represent the signals and thereby would be fed into the classifiers. Some of the used features were: mode range, first and second quartile, standard deviation, minimum, maximum rates etc. A weighted undirected network graph was used to present the electroencephalogram data whose edge weights were natural numbers. The graphs used in the study had connections between edges and nodes which were dependent on the Euclidean distance between the points. Connections that were found to be lower than a predefined threshold were eliminated from the graph. It was found that different signals were spotted and revealed using the proposed complex network graph. For the evaluation of the above methods, the least squares Support Vector Machine (LS-SVM) classifier was implemented for the final classification of the signals and several other experiments conducted separately for the evaluation of the features. It was observed that the Radial Basis Function as kernel for the LS-SVM provided the best results (accuracy 98%) comparing to all other kernels. Finally, it was shown that the use of a combination of network features gave better classification accuracy results.

In (Suryannarayana Chandaka, 2009) the authors proposed a pattern recognition technique called cross-correlation and used it with the Support Vector Machine classifier which is a common one. However, the interesting in that study was the usage of cross-correlation for the extraction of useful features which were proven to be efficient and useful concerning the accuracy rates in the classification experiments. The testing of the classifiers and the schemes were performed in order to solve the binary classification problem. Since several classifiers were tested, all of them were fed with features extracted from the cross-correlogram and generally provided robust predictions on the class of the signals. The best accuracy for the specific problem was obtained after a lot of experiments and different classifiers' implementation and it was 95.96%.

In this step, it is important to mention a few things about the cross-correlation sequence. The latter sequence measures the similarity between two signals. A matrix is used to discard useless features and provide the classifiers with the best and most important features. In the paper mentioned above, two experiments were conducted, and the features used in both the experiments were: peak value (PV), instant of peak (IP), centroid(C), equivalent width (EW) and mean square abscissa (MSA). The results of the proposed scheme were good but comparing to other studies and schemes like the one by (Mohammed Diykh, 2017) it seemed that the above method was not the best for the classification. This is an important note since we know that the two papers used the same classifiers but totally different pre-processing methods and feature extraction techniques. This fact is proof that the used features and the preprocessing section generally, play a determining role on the quality of the final predictions.

In (U. Rajendra Acharya, 2017) a 13-layer deep convolutional neural network was deployed for the classification of ictal, inter-ictal and normal EEG signals. The deep neural network was implemented without any feature extraction or selection process. However, some normalization of the data was implemented in order to enhance the efficacy of the implemented classifier. The experiments conducted using the Convolutional Neural Network classifier used two different activation functions, the linear rectified activation unit, and the SoftMax. The classifier used also weights and biases of previous layers to produce the final outputs. In addition, back propagation was deployed to train the CNN. The whole classification process was constantly updated and thereby produced better results. Noteworthy is the fact that with no feature elimination or feature extraction, the classifier achieved accuracy 88.67% which is pretty good considering that there was no other pre-processing or cleansing of the data.

One more study praising the SVM was by (Ralph Meier, 2008) and it proposed a generic and real-time automatic detection scheme of ictal patterns. Seven features were used for the classification of the EEG signals and those features were optimized using a two-step optimization method. The features potential to discriminate signals was tested separately in the first step and a self-organizing map was computed in the second step in order to inspect possible clusters. The SVM classifier was implemented in order to allocate signals into one of the two classes and the classification accuracies received from the conducted experiments were between 85% and 100%.

Another application of Support Vector Machine classifier was presented in (Andrew B. Gardner, 2006). The authors of the paper used a technique that mapped the EEG's time series signals into sequences by using some classifying statistics that were computed from one-second data frames. The SVM classifier was trained on epochs of normal and seizure data. The ictal signals presented changes in feature space and increased the empirical outlier fraction as mentioned in (Andrew B. Gardner, 2006). Then a hypothesis test took place determining if the changes in the parameters differed from their nominal values and in that way signaling seizure detection events. The major advantage and innovation of the proposed method was that the whole scheme was based on the detection of changing points in the empirical outlier fraction and that was all done with respect to a feature space which discriminated healthy from ictal signals. Furthermore, the proposed technique overcame three major problems that other algorithms failed to solve. First, the whole process did not need training and testing phase and it could be implemented in real time. Second, there was no need for patient tuning and last but not least, the method worked without knowing the consciousness state of the patient that was monitored by the EEG. (Andrew B. Gardner, 2006)

An innovative method called Hilbert Vibration Decomposition (HVD) was used for the classification of ictal and healthy EEG signals in (Mutlu, 2018). The method presented in that paper employed components with the highest energy as features and used the least squares support vector machine for the recognition of the signals. Results taken from the tests in the dataset provided 97.66% prediction accuracy. To evaluate the performance of the implemented method, the HVD feature extraction method was compared to the following schemes: Hilbert-Huang Transform, sort-time Fast Fourier Transform, and Reduced Inference distribution. Finally, it was observed that the model used in (Mutlu, 2018) decomposed EEG signals into mono-components by using the HVD decomposition method, extracted the significant features and at last, fed them into the LS-SVM classifier. Worth noted is that the aforementioned HVD Decomposition was considered an innovative and efficient method to approach the EEG signals classification problem since it was never used before in problems like seizure detection.

A study which proposed a group of algorithms and not a specific algorithmic solution was presented by (Zamir, 2016). In that paper, linear least squares preprocessing was used for better feature extraction and optimization of the accuracies to be produced by the classifiers. The EEG signals were approximated as sine waves and their amplitude was modeled as a polynomial of increased degree and a spline. Furthermore, two extraction methods were used to extract key features. Several tests on the data had been done using different classifiers and the results were evaluated using some of the basic statistical measures like sensitivity, specificity, error rate, false positive rate per hour etc. Finally, the author, observed that the best accuracy reached 100% and it was achieved by several combinations of classifiers (Logistic, Lazy IB1, Lazy IB5, and J48) and feature extraction models.

A study by (Shivnarayan Patidar, 2017) proposed and presented a novel method for classification of signals using the tunable-Q wavelet transform (TQWT) which encapsulates the sparsity of the signals and also their non-stationarity. The TQWT decomposed the EEG signals into separate bands helping in that way the feature extraction process. After the decomposition, the Kraskov entropy was computed for the different sub-bands and was used as a feature that enhanced classification. Concerning the classifier, the least squares Support Vector Machine was implemented, and it was observed that the Kraskov entropy-based feature had higher rates when the signal was ictal and lower rates when non-ictal.

An expectation maximization algorithm was introduced by the author in (Subasi, 2007) for the classification of seizures in EEG signals. An alternative of neural network classifier was used, called mixture of experts. In the paper was mentioned that the classification procedure included several steps. One of them was the decomposition of the signals into different bandwidths using the discrete wavelet transform. Afterwards, these sub-bands were used in the mixture of expert's algorithm in order to allocate the signals into one of the two classes. In an effort to increase the accuracy of the predictions, a gating function was used which provided weights to the outputs of each expert. The performance of the scheme was evaluated and compared to others and it was also observed that the accuracy achieved by the mixture of expert's model was slightly better (94.5%) than the one received from the stand-alone neural network (93.2%).

Time and frequency analysis were performed on the electroencephalogram signals dataset in the paper (Zakareya Lasefr, 2017). The Chebyshev filter was used for the preprocessing of the data by eliminating useless artifacts on the EEG signals. After that, Wavelet Analysis was deployed for the decomposition of the signals into sub-bands. From the 5 sub-bands only one was created, the Delta sub-band, which was selected by the authors and was used in the paper. For the extraction of features, Discrete Wavelet Transform was used and then a thresholding technique was applied in order to detect noise in the signal. The features used in the paper were: domain frequency of the wavelet and of the threshold and wavelet of the 5^{th} sub-band. Finally, the Artificial Neural Network (ANN) and the Support Vector Machine (SVM) classifiers were used to identify ictal and non-ictal signals. Another field relevant classifier used along with the ANN was the Feed Forward Neural Network (FFNN) which consisted of a number of neurons and several network layers. After experimenting with the classifiers and the features it was observed that the best accuracies received were 96% from the SVM and 98% from the ANN with the FFNN.

In the paper (Md. Mamun or Rashid, 2017) the authors presented another aspect of classification based completely on the features. Wavelet Analysis methods were deployed in order to extract features and several other statistical measures like sensitivity, specificity, error rate, false positive rate, and some entropies were used to the classification. The Neural Network classifier was deployed on a 2-class and 3-class classification problems. The accuracies received from the classifier were 99.5% for the 2-class classification and 79% for the 3-class problem. However, the important thing of this study was that the accuracies obtained by the classifier were dependent on the variations and changes in the number of features used in each experiment.

A new method for classification of non-stationary signals using the Optimum Allocation Technique (OAT) was proposed by the authors of (Enamul Kabir, 2016). The OAT technique was implemented for the selection of the most significant and representative signals which represented the whole dataset best. From the chosen signals some features were extracted and along with some statistical features they were loaded into the Logistic Model Tree for classification. After several experiments on the dataset and different classifiers, it was observed that the Logistic Model Tree along with the OAT gave the best accuracy performance when compared to other state of the art algorithms.

In the work of (Zeynab Mohammadpoory, 2017) a new method in which the EEG signals are mapped into a weighted visible graph (WVG) and the entropies of the graph are calculated, was proposed. In this study, the signals features were collected and loaded into 4 classifiers, chosen by the authors, in order to help classify the signals into three classes, namely ictal, healthy and interictal. The classifiers used in the experiments were Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB) and Decision Tree (DT). From the results, it was shown that the best classifier in terms of accuracy along with other statistical measures like specificity, sensitivity and error rate was the decision tree classifier with accuracy of 97%. The specific accuracy result was considered pretty good since it was obtained by the classifier with the implementation of only the following features: standard deviation, mean, and minimum of the WVG entropies.

On an innovative and highly cited article by (Alexandros T. Tzallas, 2009) a time-frequency analysis of EEG signals was presented, and several time-frequency methods were deployed and then compared to each other. The Short Time Fourier Transform (STFT) and some time-frequency distributions were used to calculate the power spectrum density (PSD) of each signal segment. In the paper, the whole signal analysis was performed in three main steps. First, the time-frequency analysis of the power spectrum density of every signal segment, second the extraction of features and measurement of the energy for each time-frequency segment and finally the classification of the segments using the Artificial Neural Network classifier. The classifier's hidden units were the sigmoid along with hyperbolic tangent for activation function. Four other classifiers were tested in the paper (Alexandros T. Tzallas, 2009) and the results of their classification helped identify the best fit classifier for the discrimination of ictal and non-ictal signals. From them, the best classifier was proven to be the initial choice, the ANN with accuracy 99.8%

Another model used for signal classification is the expert model along with the Probabilistic Neural Network classifier as it was implemented in the (Tapan Gandhi, 2010) paper. The process described in the article was based on Discrete Wavelet Transform for the extraction of features and estimation of energy in each tree decomposition node. After the formation of ictal and healthy epochs, the epochs were decomposed into detail and approximation coefficients. Features like energy, entropy and standard deviation extracted by using the discrete wavelet transform, were fed to the Probabilistic Neural Network and the accuracy achieved by the scheme was 99.33%

In the study of (Tao Zhang, 2018) a new method synthesizing Generalized Stockwell Transform with Singular Value Decomposition and Random Forest as classifier was presented. The utilization of generalized Stockwell transform used the available data and developed a time-frequency matrix. The values of the matrix were extracted using the singular value decomposition technique. Then the value vectors were fed into a Random Forest classifier and its performance was measured. The proposed scheme managed an accuracy rate of 99.63% the highest.

A major problem when processing time series is that they often have features of different time scale. In order to battle this obstacle, an end-to-end Multiscale Convolutional Neural Network (MCNN) was proposed in the paper of (Zhicheng Cui, 2016). The MCNN model combines feature extraction and classification in one framework. Using a multiple branch layer and some learnable convolutional layers, the MCNN was able to extract features with different characteristics and scales, thing that led to better feature representation. (Zhicheng Cui, 2016) The network was tested and among others, it was found that it provided the best results in terms of prediction accuracy. Finally, the MLNN was considered as a time effective and efficient method for classification of time series signals in the case of epilepsy detection.

In (Lorena Orosco, 2016) a pattern recognition neural network (PRNN) with Bayesian regulation was selected in order to classify the EEG signals and predict ictal and non-ictal events. The aforementioned method was applied to a total of 275,048 segments of 2s of duration. 3267 of the seizure segments corresponded to 18 pediatric patients with intractable epileptic seizures. The study by (Lorena Orosco, 2016) proposed an innovative algorithm for offline seizure detection. The data was gathered from 18 *pediatric* patients and 2-classification methods were tested. The first one was the Linear Discriminant Analysis and the second was the common Neural Networks. The Linear Discriminant Analysis is a combination of discriminant features that help the maximization between different groups and achieve minimization in differences within the group. After trials, it was shown that the best statistical measures were achieved using the linear discriminant analysis namely, sensitivity of 99.9 % and false positive rate per hour at 0.3. Thereby, the classifier was proven to be reliable and more robust than the Neural Network.

In a different from the already examined literature paper, the authors of (Emina Alickovic, 2018) created and tested several models in order to predict the *onset* of seizures in patients and finally created a fully automated and functional method. Although the main concern of the paper was to predict seizure onset, there were several experiments conducted, with purpose the simple predictions and seizure identification. The whole process contained data from two different databases and four main steps were followed. Initially, for the denoising of the signals, principal component analysis was implemented. In the second step, the signals were decomposed using either empirical mode decomposition, wavelet packet decomposition or discrete wavelet transform and then using some statistical measures, the authors extracted useful features that in the final step were used by the machine learning algorithms. Some of the classifiers implemented in several pairs with different feature extraction methods were: Random Forest and SVM. The models achieved good accuracies on predicting the ictal from the healthy signals (99.7%) as well as good accuracy rates for the prediction of seizure onset.

The authors of (Md. Kamrul Hasan, 2017) implemented the commonly used K-Nearest Neighbor classifier, and several entropies and statistical measures to predict the ictal and non-ictal signals from patient’s electroencephalograms. The features used in the paper were: standard deviation, standard error, approximate entropy, modified mean absolute value, zero crossing, and roll-off which were all extracted from epileptic signal. These features were then used by the K-NN algorithm which produced a prediction for each signal segment. It should be mentioned that all the statistical measurements were taken by ictal signals and thereby represent only one class of the data. Furthermore, noteworthy is the fact that the above process was mainly based on the statistical measures and the paper’s purpose was to show that the procedure of the prediction was done by using a statistical approach and less an algorithmic method.

In contrast to the previous study, the authors of (Ramy Hussein, 2018) presented a robust algorithmic method for seizure detection which was proven to work well in real-time data and had good generalization rate. In the proposed method, a feature cleansing method was developed and applied to the signals called L1 penalized robust regression which spotted the most important features for the classification. The unimportant ones were discarded, and the useful ones were used by the K-NN classifier. The results have shown that the accuracy from the proposed model reached 100% which was a surprise since the specific algorithm has already been used in other schemes but only this method gave 100% accuracy. Finally, after a lot of research, the authors noted that the algorithm was proven robust and useful even when the data contained noise and recurring useless and misleading patterns.

2.4.2 Classification processes

A lot of algorithmic models and techniques have been implemented into various datasets for the prediction of existence or absence of seizures. Some of the most used models for the seizure prediction were found to be Support Vector Machine, Artificial Neural Networks, Random Forest and

Decision Trees. In the majority of the examined papers, the authors implemented at least two different classifiers. In Fig. 3 it can be seen that the most commonly used classifier is the SVM.

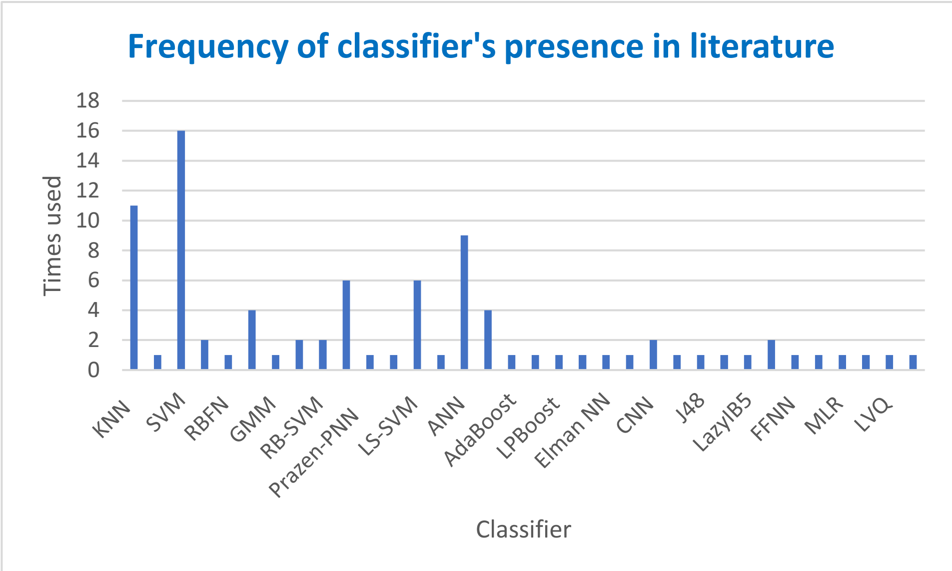

Fig.3

Also, for the creation of significant feature spaces, several methods were implemented and some of them were Discrete Wavelet Transform, Continuous Wavelet Transform, Fast Fourier Transform, Random Forest, power spectral analysis, time-frequency models, t-distributed Stochastic Neighbor Embedding etc. In the studies presented above two kinds of classification problems appear. The most common one is the 2-class classification problem in which the goal is to classify ictal from healthy EEG signals and the other one is a 3-class classification problem. In the latter, the purpose is to allocate the signals to healthy, ictal or inter-ictal class.

Apart from the different classification problems, there were 2 different scopes observed in the examined literature. Most of the cases had as purpose to create models and enhance or create an automation process for seizure classification. In the first group of studies, the purpose was to create a model good for real-time classification of seizures which would provide robust prediction and be easily usable. The other group concentrated more on finding models and testing classification schemes for best predictions. In that case, the purpose was to have very good predictions in terms of accuracy, but a drawback was that the predictions in most cases were found to be data dependent.

2.5 Results and Synopsis of the examined literature

The examined literature has shown that there are many different approaches concerning the classification and feature extraction scheme to be chosen for seizure detection problems. In most of the cases, the classification problem was to discriminate EEG signals to one of two classes: ictal and healthy. However, there were some cases in which the classification problem was to allocate the signals into ictal, inter-ictal and healthy EEG signal. It was observed that the majority of the studies of the literature review used data from both healthy and sick adult individuals but yet there were a few cases in which the EEG signals were recorded from pediatric patients.

Concerning the basic part of classification of the signals in most studies, it was stressed the fact that the feature extraction process was very important and contributing into the final classification accuracy. Most of the times the key features were the ones that made the difference when loaded and tested into different algorithms. As mentioned before, in almost all papers more than one classification schemes were proposed and tested. As a result, there are many different algorithmic models that fit and solve the classification problems and their results were evaluated with a wide range of evaluation metrics and statistical measures. However, the most used metrics were the accuracy, sensitivity, specificity and error rate. From a closer look at the classification accuracies and the results of all the studied schemes, it was spotted that the range of values that the models produced were between 65% and 100%. Those rates though do not represent the most frequently acquired accuracy since they are border rates. The average accuracy was around 95% which is considered a very good rate for classification especially if we consider that many studies proposed real-time classification schemes and others dealt with 3 classes instead of two. Finally, noteworthy is the fact that all studies that dealt with a three-class classification problem presented considerably lower accuracies than those produced by the two-class classification papers.

3. The Dataset and Visualization

In this part of the thesis, a detailed explanation of the dataset used in the study is displayed. Visualization and tables were used in order to help understand better the data and their meaning which is considered the most important step in the whole research process.

3.1 Data selection and pre-processing

The dataset used in this thesis was taken from University of Bonn, Germany and the Department of Epileptology. The initial dataset consisted of 5 folders each containing 100 files. Each file contained 23,6 seconds of brain activity recording which corresponded to a different person. The time series signals were sampled into 4097 data points. Each of the 4097 data points corresponded to the value of the EEG recording at a different time. So, there are 500 individuals/files, each sampled in 4097 data points and each of 23,6 seconds duration. In this work, 3 folders with each containing 100 files were used. Two of the folders contain non-seizure signal and one contains seizure signals. File F contains EEG signals from individuals that used to suffer from seizure. Signals were recorded after the conduction of a surgical procedure in which the hippocampal formation responsible for seizures on the epileptogenic zone was resected. Thereby, the patients whose EEG signals were recorded were in a controlled seizure state and some of them were "cured". S file contains non-seizure signals taken from healthy patients who were in semi-wakefulness state during the EEG having their eyes closed. Finally, the O dataset contains seizure signals coming from people suffering from the disorder while the signal was recorded during a seizure. Z and the N file contained non-ictal data. Those files though were not used in the study since it was considered that they did not add any useful information on the inquiry of the thesis.

The data were stored in 3 different excel files in which there were 100 columns containing the information of the 100 individuals and 4097 rows each containing a value of each different data point of the sampling frequency. An example of the initial structure of the dataset is given in the figure below in which a part of the F dataset is shown.

F001	F002	F003	F004	F005	F006	F007	F008	F009	F010
34	60	26	26	13	-15	-24	23	-263	59
33	47	16	16	6	-2	-27	17	-263	52
28	38	13	13	-1	0	-23	10	-261	51
22	29	12	12	-13	2	-28	10	-258	46
21	28	17	17	-29	-2	-34	7	-258	43
22	30	16	16	-42	-1	-40	-5	-251	43
22	28	7	7	-53	-2	-47	-6	-241	36
19	30	-6	-6	-71	-7	-43	-13	-236	26

Fig.4

In Fig. 4 the first ten columns of the F dataset and information on the first 8 data points/rows are displayed. The values of the table represent electric current measurements taken from the EEG after they were imposed into a sampling process. Before moving on to the visualization, an identification column was added in each of the three folders. The column was named "ID" in all files and it contained the number of the row. That column addition was performed because after testing it was found that the visualization results and plots were more comprehensible. Finally, the 3 folders have 101 columns with the first being the ID and 4097 rows.

3.2. Visualization and plots

The visualization plots to follow present all the differences and similarities between the different signals from both the same and different folders. The major differences as it will be proven were found to be between ictal and non-ictal signals. The signals are presented using the same scale in the axes for better visualization and perception of differences in the signals. The x-axis represents the column visualized each time and the y-axis represents the rows/ data points of the measurements.

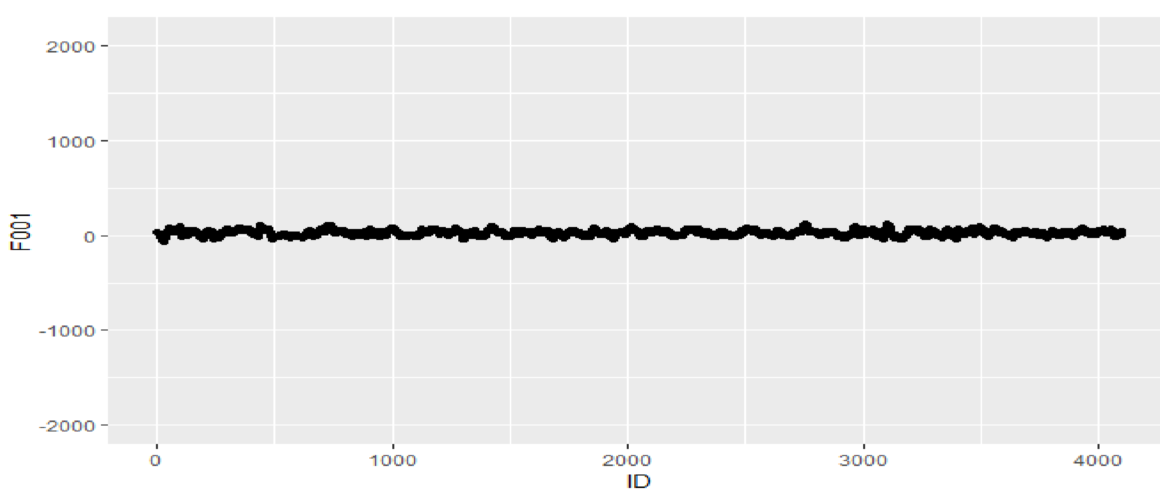

Fig.5

In Fig. 5 the signal does not present extreme spikes and the frequencies that the signal fluctuates are limited. From further study, it was found that in the F file all the frequencies were between a range -1147 and 2047. This range of values was considered wide and during the research of the signals it was found that most of the recordings in the F file fluctuated in a smaller range of frequencies, usually between -500 to 500.

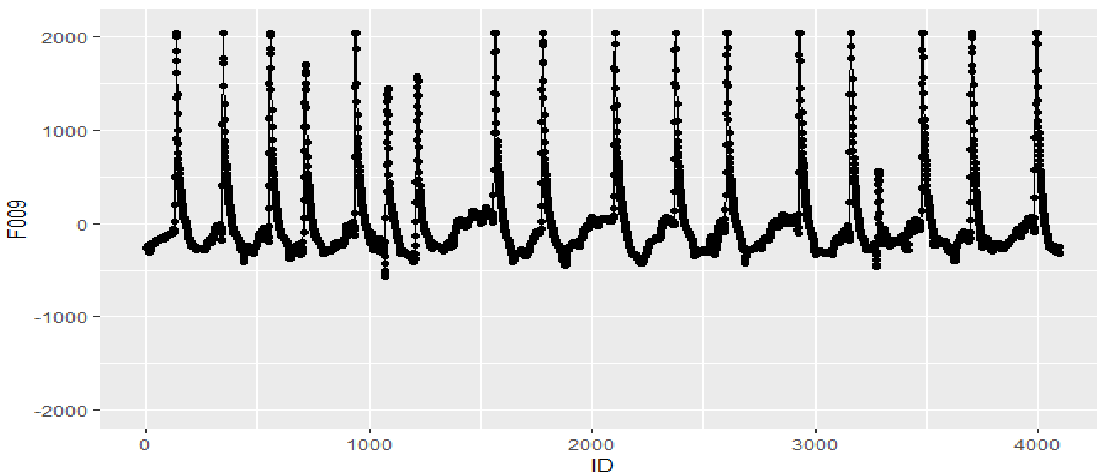

Fig.6

Apart from the above smoothed and homogeneous recording there were found signals as the one presented in Fig. 6 which presented unexpected spikes. From further research on the signal from people that suffered from the disorder in the past, it was found the latter signals mostly contained spikes of similar amplitude which were equally distributed during the recording.

Though, as it can be seen in Fig. 7 there were some recordings which presented intense fluctuations with many spikes distributed randomly within the recording. Those recordings were limited but they consist useful examples, worth studying since they present anomalies not usually seen in this group of signals.

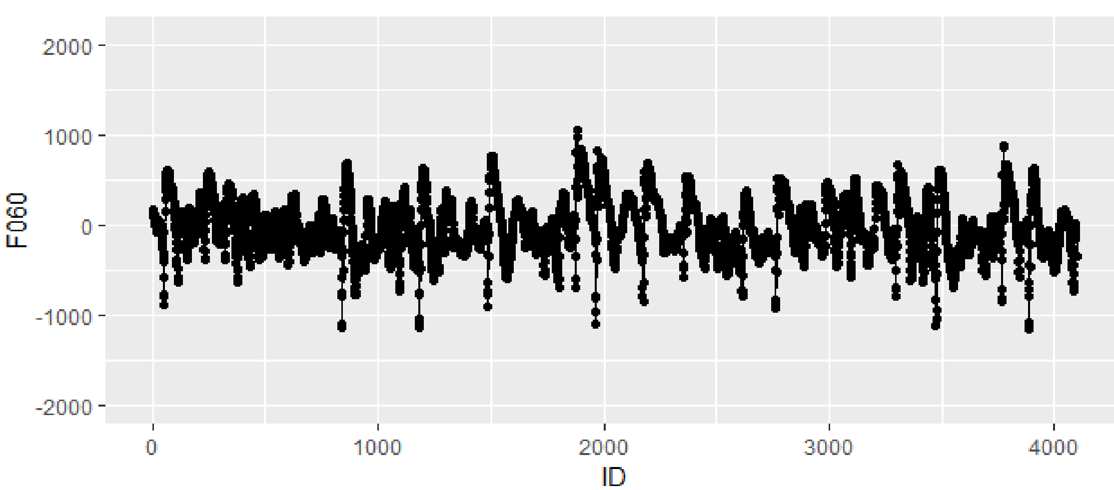

Fig.7

The O dataset contains signal coming from completely healthy people that took the EEG while awake in a relaxed position with their eyes open. This dataset was selected because it was considered interesting to find the correlation and similarities between the signals of healthy and in active consciousness state individuals with the signals in the other two categories. A usual EEG recording of a healthy person is the one seen in Fig. 8

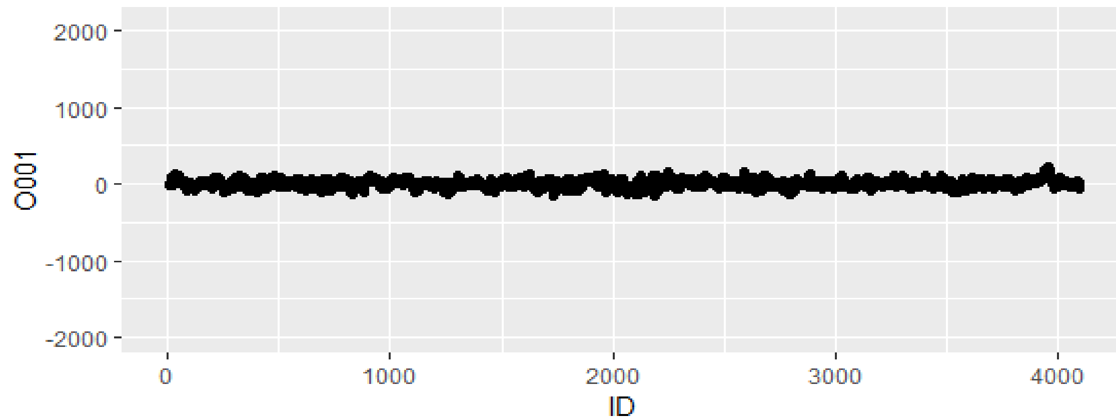

Fig. 8

It can be observed that the signal is fluctuating in a small range of values and is condensed comparing to the signals from the F file which is more evenly distributed. Namely, the signals in folder O were between -427 and 360 but in most cases, the range of most recordings was much smaller. Apart from the mainly observed signals there were spotted recordings with a different appearance in terms of spikes and range frequency. In the following figure, Fig. 9 the signal is fluctuating in a wider range of frequencies and a few spikes are observed. It should be mentioned that the spikes recorded here are not as high as those spotted in the data from people who used to suffer from epilepsy. The signals recorded from healthy individuals seem to have a smooth and more stable expression and of course much less spike amplitude comparing to the F dataset's spikes which were intense.

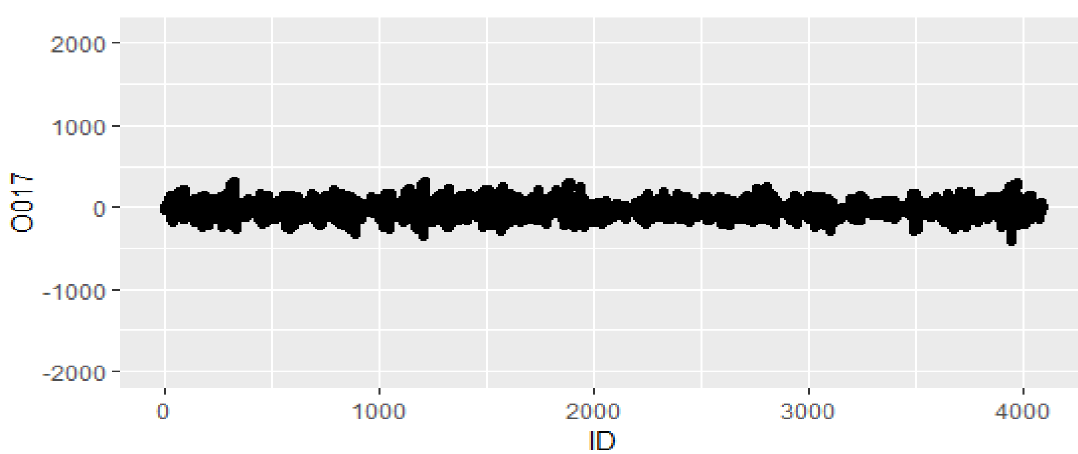

Fig. 9

The final dataset selected and used in this thesis was the S dataset containing the seizure signal. The recordings of this folder are very different from those of the other two folders. As seen below in Fig. 10 the signals are much more intense and fluctuation in a wide range of values is observed. There are several spikes with different amplitude and with no obvious patterns. Worth mentioning is the fact that as many spikes in the vertical axis are above zero roughly speaking that many are below zero.

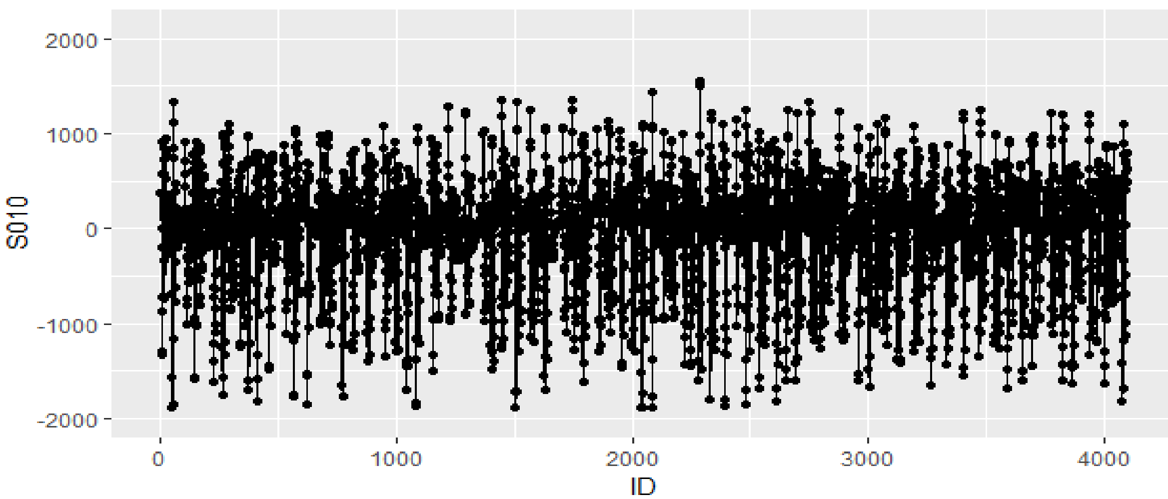

Fig. 10

Though there are cases like the one observed in Fig. 11 in which the spikes appear more in the positive axis while in the negative axis the signal is more conventional and without big changes in the amplitude. The range of values observed usually in seizure signals was found to be between -1885 and 1793 and it should be mentioned that most of the recordings fluctuate around those values and so these values are not some extreme rates as it happened in the other two datasets.

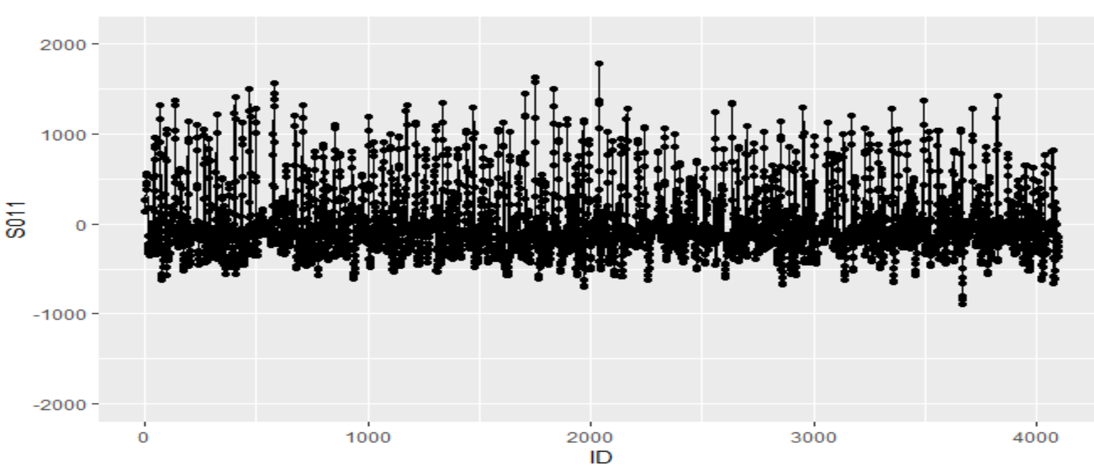

Fig. 11

For further exploration of the data, correlation matrices were created. After a detailed inspection of the matrices, it was found that there was no obvious correlation between the signals of the same folder. An example is given in the figure below Fig. 12 and it shows that there is no correlation neither between the same folder signals nor within different folder recordings.

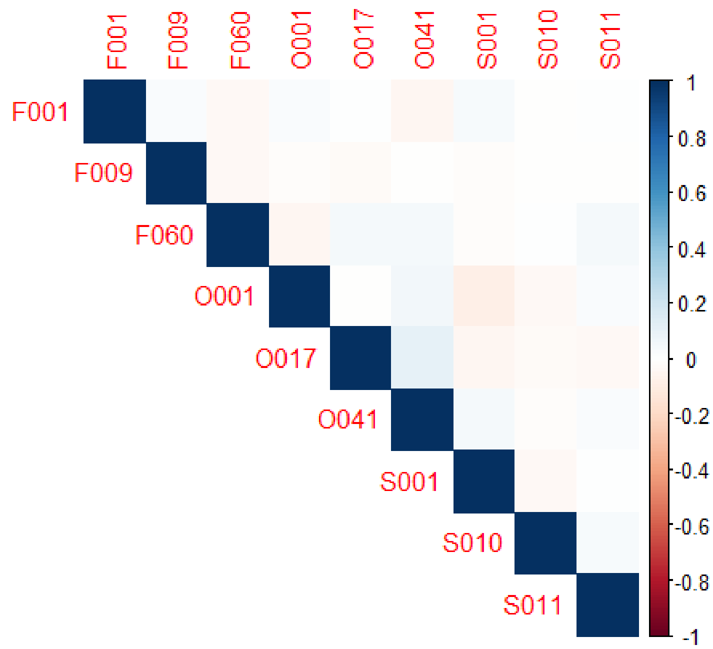

Fig.12

It can be seen that all squares have very faint colors meaning that the correlation between them is very weak. That fact was expected since EEG signals are non-linear and non-stationary and brain spikes are not formed due to specific patterns or triggers of the brain, especially on the seizure signal. So, from the correlation matrix, it was shown that there was no clear relation between the signals and it was inferred that a signal does not have any impact on others. That fact was confirmed also in the literature where it was stated that signals from different individuals are never the same. Specifically, it was mentioned that there are as many different ways to inspect and interpret signals as many EEG recordings there are. From that statement, partial independence between the recordings was considered a fact.

4. Algorithmic modeling

The main scope of this chapter is to explore the algorithms and models used to describe and accurately predict seizures. In order to find the most useful and effective algorithms, many trials were performed on the dataset using Weka 3.8 data mining tool. After finding the algorithmic model fitting the data in the best way, a short description of each of the algorithms was presented. Finally, in the last part of this chapter, the results of the classifiers were displayed, after the default options had been changed.

4.1 Preprocess of the data

Before loading the data to the program some processing steps and changes on the files were performed. Initially, as mentioned before the data files contained data of 100 individuals each, which were displayed into 100 columns in each excel file. For the sake of efficiency, all three data tables were transposed so that rows became columns and columns became rows. The result was that the files after that process contained 100 rows and 4097 columns and in that way, each row represented an individual and contained information of 23.6 seconds which was "encoded" into the 4097 data points displayed in the file. In each one of those files, an extra column was added which contained the class label in which the signals belong. The class labels for the sake of better visualization and process in weka were allocated as follows: "h" for healthy signals and "s" for seizure signals. After that, the three separate files were merged into one excel file which contained 301 rows with the first being the enumeration line of the 4097 data points and 4098 columns with the last one containing the class label of each individual as it can be seen in Fig. 13.

4091	4092	4093	4094	4095	4096	4097	class
22	31	40	45	39	41	7	h
919	916	829	722	512	130	196	s
86	99	113	119	114	99	-130	h
-216	-525	-735	-847	-721	-491	910	s
-195	-218	-234	-238	-209	-165	-212	s
296	248	209	177	149	126	42	h
430	472	515	527	480	397	217	s
156	156	153	150	146	140	191	s
106	104	100	98	101	99	-554	s
86	99	113	119	114	99	-130	h

Fig. 13

Figure 13 displays the ten first signals along with seven of the 4097 data points expressing each signal. The last column represents the class and contains either "s" or "h". The class attribute has values in every single individual/row, meaning that there are no missing values. Finally, the excel file was converted into a csv so that it could be usable by Weka.

4.2 Classification algorithms results and performance

After the preparation and preprocessing the dataset was loaded into Weka 3.8. The first step after loading the data into the program was to set the last attribute as class and after that, the classification algorithms were tested. Since the purpose of the thesis is to find an algorithmic model that fits best the data, several tests on all the available classifiers were performed. Weka provided eight groups of algorithms and all the available classifiers were implemented. In order to compare and distinguish the best algorithms as forecasting models for epilepsy detection, the accuracy of each classifier, error rate, the time taken to build each model, and other available measures and statistics were recorded. All the available classifiers were implemented using their defaults options and 10-fold cross-validation.

The first group of algorithms tested in weka was the Bayes which used the Naïve Bayes classifier in several versions and produced prediction results. The implemented algorithms from this group were four.

- Bayes Net model was built in 1.56 seconds and gave 93.6667 % classification accuracy. Kappa statistic which represents the agreement between the predicted and the actual class was found to be 0.8542 and the F-statistic was 0.936. From the confusion matrix and visualization of the errors, it was found that the algorithm misclassified 14 sick individuals as healthy and 5 healthy as sick.
- Naïve Bayes model gave 93.3333 % accuracy, was built in 0.55 seconds with 0.8477 kappa statistic and weighted average F measure at 0.933. Here the misclassified instances were 20 in total and 7 of them were classified as sick while healthy and the rest as healthy while sick.
- Naïve Bayes Multinomial Text classifier gave the lowest accuracy from the Bayes group which was found 66.6667%. The model was built in 0.01 seconds and presented kappa statistic rate 0 which means that there was no agreement between the predicted outcomes and the actual ones. The important in this classifier was that all the instances were classified as healthy. So, 100 instances out of the 300 were misclassified.
- Naïve Bayes Updateable produced the same accuracy with the Naïve Bayes which was 93.3333% and the model was built in 0.56 seconds. Kappa statistic was found to be 0.8477, average weighted F-score 0.933 and the misclassified instances were also 20 with the same error distribution as in the Naïve Bayes.

The figures below, display the visualization of the errors from Bayes Net and Naïve Bayes Multinomial Text models respectively. The small squares represent the misclassified instances in every case and the blue color is for healthy class and red the sick/seizure class.

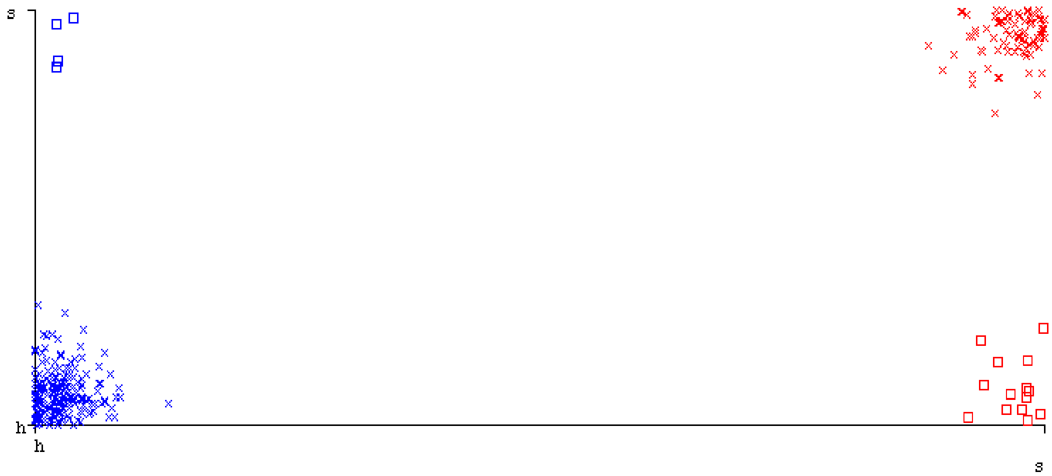

Fig. 14

As expected in Figure 14 some squares are spotted on the left upper corner and some on the right bottom corner showing the misclassified instances of the Naïve Bayes model. On the other hand, in Fig. 15 it is obvious that there are a lot of misclassified instances and they are all of one kind. In this case, there are all the sick people who were mislabeled as healthy.

Fig. 15

Next group of classifiers implemented was the lazy. This group contained algorithms based on the lazy learning process. The implemented classifiers were the following: IBk which implements the K nearest neighbor classifier, KStar, and LWL (locally weighted learning).

- KStar took 0 seconds to train the model and gave accuracy rate of 66.6667 % with kappa statistic at 0. KStar algorithm as the Naïve Bayes Multinomial Text presented the same confusion matrix having all the seizure instances presented falsely as healthy.

- LWL (locally weighted learning) gave 75% classification accuracy and needed 0 seconds to build the model. The kappa statistic was 0.3363 and the average weighted F measure 0.712. The confusion matrix presented a lot of misclassified instances. 75 in total with 68 being seizure signal classified as healthy and 7 instances where the signal was healthy, and it was classified by the algorithm as sick.
- IBk achieved 81.6667 % accuracy and 0.01 seconds were used for the training of the model. The kappa statistic was found to be 0.5217 and the average weighted F-score of 0.793. From the confusion matrix, it was observed that there were several mislabeled sick instances which were classified as healthy. In Figure Fig.16 below, it can be seen that the performance of the model is poor when it comes to classifying seizure signal while at the same time the algorithm makes no mistakes at accurately labeling the healthy EEG recording signals.

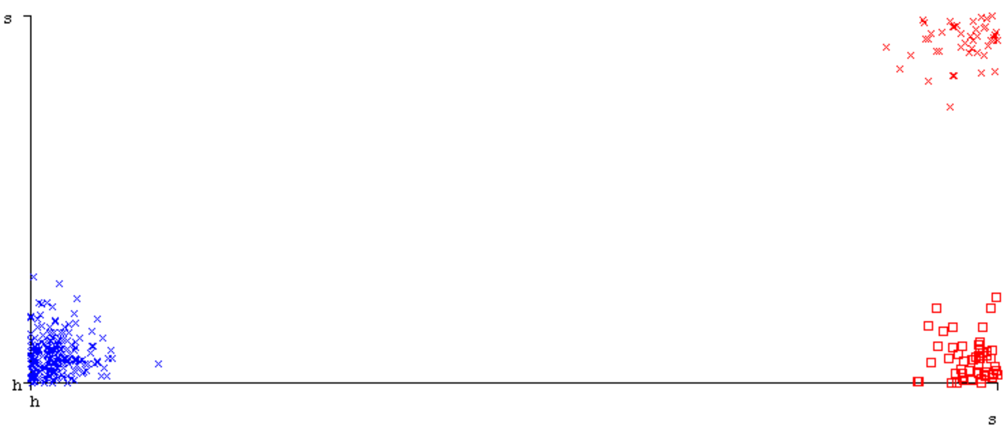

Fig. 16

From the group of meta algorithms most of the available classifiers were tested. Though one and specifically multilayer perceptron was omitted, due to high computational demands and not enough computational power of the computer.

- AdaBoostM1 model took 6.4 seconds to build and achieved 86.3333% accuracy. The F- score was 0.855 and the kappa statistic 0.6649. The misclassified instances were 41 in total and most of them were sick signals mislabeled as healthy.
- Attribute Selected Classifier gave 89% classification accuracy and the model needed 461.18 seconds to get build. The kappa statistic was 0.7519 and the F score 0.890. In this model, it was observed that the misclassified patterns were evenly distributed. Specifically, there were 17 cases of classified as healthy but actually sick individuals and 16 cases of sick individuals falsely classified as healthy.
- The Bagging classifier managed 90.6667% of accuracy and the model was built in 7.19 seconds. The F measure was 0.903 and the kappa statistic 0.7778. The misclassified instances were 28 and 25 of them were mislabeled as healthy while having seizure events. This kind of

error is considered very important since the algorithm fails to identify seizure signals in their appearance and in that way many patients could be falsely diagnosed. In the figure Fig. 17 below the distribution of the errors are presented.

Fig. 17

- Classification Via Regression algorithm managed 81.3333 % accuracy score accompanied by F-measure 0.805 and kappa statistic 0.5508. Out of the 300 individuals, 56 were misclassified with 41 of them labeled as healthy while sick.
- CV (cross-validation) Parameter Selection, Vote and Weighted-Instances-Handler-Wrapper algorithms were built in 0 seconds and achieved 66.6667% accuracy with kappa statistic rate at 0 and 100 misclassified as healthy instances.
- Filtered Classifier managed 89.6667% accuracy and took 1.38 second to build. The F measure was found to be 0.897 and kappa statistic was 0.7726. The misclassified instances were 31 and 20 of them were labeled as sick while they were actually healthy.
- Iterative Classifier Optimizer algorithm gave accuracy 87.6667% and kappa statistic 0.7071. The time taken to build the model was 41.63 seconds and the F score was found to be 0.873. From the misclassified recordings 29 were labeled as healthy while sick and 8 were labeled sick while healthy.
- LogitBoost achieved 88.6667% accuracy with 0.7316 kappa statistic and F measure with rate 0.883. The time needed to build the model was 2.79 seconds and from the confusion matrix in the figure below Fig.18, it was observed that there were 34 misclassified signals and 27 of them were labeled as healthy while they were actually ictal.

LogitBoost	classified as	
actual class	healthy	sick
healthy	193	7
sick	27	73

Fig. 18

- The time taken to build the Multi-Class-Classifier model was 499.02 seconds and the accuracy achieved was 77.6667%. The kappa statistic was found to be 0.4274 and the average weighted F-measure 0.753. The misclassified signals were 67 in total and 58 of them were ictal signals mislabeled as healthy.
- Multi-Class-Classifier-Updateable_achieved a better accuracy than the simple Multi Class Updateable and it was found to be 83%. The model was built in considerably less time since it took 2.78 seconds to build. The F score was 0.815 and kappa statistic 0.5714. This model misclassified 47 ictal signals as healthy and 4 healthy signals as seizure.
- The Multi Scheme model was built in 0.02 second achieving 66.6667% accuracy. The kappa statistic as expected was 0 and all the 100 ictal signals were misclassified as healthy. The model identified only healthy recordings and ignored all the ictal cases.
- Random Committee classifier was built on training data in 0.31 seconds and managed an accuracy of 94.6667 %. The kappa statistic was found to be 0.8763 and the average weighted F measure 0.946. From the 300 instances, 16 were misclassified by the model and 14 of them were classified as healthy while they represented ictal signal. In Figure 19 it can be seen that there are only a few misclassified instances.

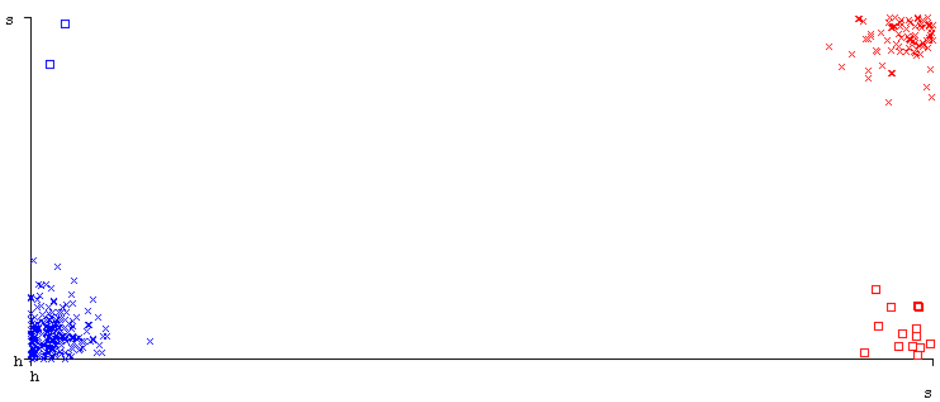

Fig. 19

- Randomizable-Filtered-Classifier achieved 86.3333% accuracy and was built in 0.17 seconds. Kappa statistic was found to be 0.6612 and F-score 0.854. From the confusion matrix below, it is obvious that most of the misclassified instances were ictal, but they were falsely labeled as healthy.

Randomizable Filtered Classifier	classified as	
actual class	healthy	sick
healthy	198	2
sick	39	61

Confusion matrix with the secondary diagonal showing the misclassified instances.

- Random Sub Space classifier was built in 4.11 seconds and achieved 93.6667% accuracy accompanied by 0.8527 kappa statistic and 0.935 F measure. The misclassified instance was 19 in total and only 3 of them were classified as sick while they were healthy.
- Stacking model was built in 0.01 seconds and gave 66.6667% accuracy. Kappa statistic was found to be 0 and all the ictal signals were mislabeled.

A different group of classification algorithms was the rules and it contained models based on certain classification rules.

- Decision Table classifier achieved accuracy score of 87.3333% while it took 28.5 seconds to build. The kappa statistic for this model was found to be 0.7092 and F score was 0.872. Finally, the misclassified instances were spotted on the confusion matrix from which it was found that the total number of the mislabeled recordings were in total 38.
- JRip classifier was tested next and managed classification accuracy of 86% accompanied by a kappa statistic rate of 0.6912 and F-score of 0.861. The model was built in 7.25 seconds the number of misclassified instances were 42 in total. Noteworthy is the fact that while most classifiers mislabel ictal signals as healthy in this case the opposite happened. 25 instances were classified as sick while they were actually healthy, and 17 were labeled as healthy while they were ictal.
- Next model tested was the OneR which achieved 78.3333% accuracy and was built in 2.08 seconds. Kappa statistic was found to be 0.4715 and the average weighted F measure was 0.771. At last the misclassified instances were in total 65.
- PART which is a decision tree algorithm was built in 3.39 seconds, achieved 89.6667% accuracy with kappa statistic at 0.7634 and F measure 0.896. The mislabeled instances were 31 in total.
- ZeroR managed 66.6667% accuracy and was built in 0 seconds and as previously it classified correctly only the healthy instances while failed to classify correctly ictal signals.

A very useful and widely used group of algorithms is the trees group. This group contained models using the idea of decision trees in different versions.

- Decision Stump classifier achieved accuracy 75.6667% and took 0.4 seconds to be built. Kappa statistic was found to be 0.3615 and F measure 0.723. In total, the number of mislabeled signals were 73 and most of them were classified as healthy while actually ictal.
- Another model achieving a very good accuracy rate of 93% was the Hoeffding Tree which was built in 2.21 seconds. Kappa statistic rate was 0.8405 and average weighted F measure

0.930. From the misclassified instances 13 were labeled as healthy while they were ictal and 8 were labeled as sick while they were actually healthy.

- The J48 algorithm was tested next, managing 90.6667% accuracy and F score rate of 0.906. The model was built in 1.8 seconds and kappa statistic was found to be 0.7889.
- LMT (Logistic Model Tree) model was built in 62.95 seconds reaching accuracy rate of 80.333% and F-measure of 0.782. Kappa statistic rate was found to be 0.4957 and the number of misclassified instances were 59 in total.
- An important algorithm with good results was Random Forest which needed 1.43 seconds to be built. The model achieved classification accuracy score of 93.6667% and F-score 0.935. Finally, kappa statistic was found to be 0.8527. In total the misclassified instances were 19.
- Random Tree classifier achieved 88% accuracy and F score of 0.878. The model was built on the training data in 0.03 seconds and presented kappa statistic rate 0.7216. The misclassified instances were in total 34 and 2/3 of them were classified as healthy while ictal and 1/3 of them were classified as ictal while healthy.
- The last algorithm tested in this group was REPTree which was built in 1.07 seconds and achieved 81.6667%. Kappa statistic value was 0.5844 and F-score 0.816. Finally, the misclassified instances were 55 in total and 29 of them were classified as healthy while they were actually ictal.

Finally, the last group of algorithms tested on the dataset was the functions group. That group contained seven available classifiers which were all tested apart from the Multilayer Perceptron which failed to run due to high computational power demanded.

- Voted Perceptron was built in 0.33 seconds and achieved an accuracy rate of 59% accompanied by F-measure 0.602. Apart from the above, kappa statistic was found to be 0.1782 which is a considerably low rate and the number of misclassified instances was 123 in total. Important to note is that most of the mislabeled instances, precisely 86 of the 129, were actually healthy but they were classified as ictal.
- SMO classifier used the support vector machine algorithm with Platt's sequential minimal optimization. The model took 2.45 seconds to build on the training data and achieved accuracy 81.6667%. Average weighted F measure was found to be 0.793 and the kappa statistic 0.5217. Noteworthy is the fact that all the misclassified instances were labeled as healthy while they were actually ictal. Those cases were 55 in total and they are all gathered in the right lower corner on Figure 20 below**.**

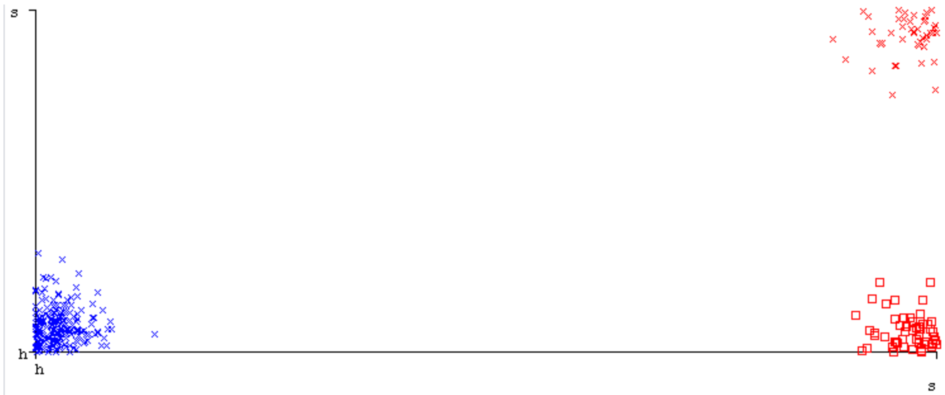

Fig. 20

- Simple Logistic classifier achieved 80% accuracy and took 15.42 seconds to build using training data. The misclassified instances were 60 and 59 of them were ictal but falsely classified as healthy. The F score of the classifier was 0.772 and the kappa statistic 0.4737. As obvious in the figure below only one instance was healthy but classified as ictal.

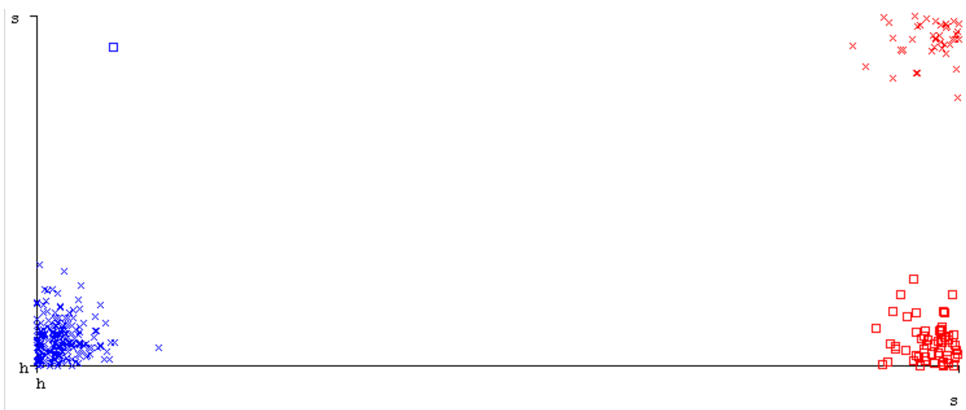

Fig. 21

- SGD Text classifier used stochastic gradient descent for learning a linear binary class Support Vector Machine. (Frank, http://weka.sourceforge.net, n.d.) The model was built in 2.22 seconds and achieved 66.6667% accuracy. Kappa statistic was 0 and all the ictal instances were classified as healthy.
- SGD is an alternative algorithm using the Support Vector Machine classifier for the learning part. The model reached 83% classification accuracy accompanied by 0.5714 kappa statistic rate and F-score 0.815. The misclassified instances were 51 in total and most of them were actually ictal but were labeled as healthy as usual.
- Logistic classifier which is based on multinomial logistic regression model was built in 286.25 seconds and achieved 77.6667 % classification accuracy. The model managed a kappa

statistic rate of 0.4274 and average weighted F-measure of 0.753. Finally, from the confusion matrix, it was observed that the number of misclassified instances were 67 in total with 58 of them being classified as healthy recordings while they were actually ictal.

At last, there were two more groups of algorithms available in weka but their results were not used. These groups were "misc" which provided two algorithms for classification and "time-series" which provided only one algorithm. Input-Mapper-Classifier and Serialized-Classifier were provided by misc group and the Serialized-Classifier did not run at all with the data, while the Input Mapper produced 66.6667% accuracy and misclassified all the seizure recordings. The time-series group provided the Holt Winters classifier which provided 13.3333% accuracy which was the lowest accuracy produced by all the tested models. Those groups of algorithms were considered not useful since they did not add any new knowledge and their performance was mediocre to totally bad.

4.3. Algorithmic models and classification results

After running all the available classifiers in weka 3.8 the results of accuracy, the time taken to build each model in seconds, kappa statistic rate, and F measure were written down in table 1 presented below.

Table 1: Models Performance

	Algorithm	Time to built(sec)	Accuracy	Kappa statistic	F-measure	Total misclassified instances
Bayes	Bayes Net	1.56	93.6667%	0.8542	0.936	19
	Naïve Bayes	0.55	93.3333%	0.8477	0.933	20
	Naïve Bayes Multinomial Text	0.01	66.6667%	0	NA	100
	Naïve Bayes Updateable	0.56	93.3333%	0.8477	0.933	20
lazy	Kstar	0	66.6667%	0	NA	100
	LWL	0	75%	0.3363	0.712	75
	IBk	0.01	81.6667%	0.5217	0.793	55
meta	AdaBoostM1	6.4	86.3333%	0.6649	0.855	41
	Attribute Selected Classifier	461.18	89%	0.7519	0.890	33
	Bagging	7.19	90.6667%	0.7778	0.903	28
	Classification Via Regression	6.33	81.3333%	0.5508	0.805	56
	CV Parameter Selection	0	66.6667%	0	NA	100
	Vote	0	66.6667%	0	NA	100
	Weighted-Instances-Handler-Wrapper	0	66.6667%	0	NA	100
	Filtered Classifier	1.38	89.6667%	0.7726	0.897	31
	Iterative Classifier Optimizer	41.63	87.6667%	0.7071	0.873	37
	LogitBoost	2.79	88.6667%	0.7316	0.883	34
	Multi-Class-Classifier	499.02	77.6667%	0.4274	0.753	67
	Multi-Class-Classifier-Updateable	2.78	83%	0.5714	0.815	51
	Multi Scheme	0.02	66.6667%	0	NA	100
	Random Committee	0.31	94.6667%	0.8763	0.946	16
	Randomizable-Filtered-Classifier	0.17	86.3333%	0.6612	0.854	41
	Random Sub Space	4.11	93.6667%	0.8527	0.935	19
	Stacking	0.01	66.6667%	0	NA	100

rules	Decision Table	28.5	87.3333%	0.7092	0.872	38
	JRip	7.25	86%	0.6912	0.861	42
	OneR	2.08	78.3333%	0.4715	0.771	65
	PART	3.39	89.6667%	0.7634	0.896	31
	ZeroR	0	66.6667%	0	NA	100
trees	Decision Stump	0.4	75.6667%	0.3615	0.723	73
	Hoeffding Tree	2.21	93%	0.8405	0.930	21
	J48	1.8	90.6667%	0.7889	0.906	28
	LMT (Logistic Model Tree)	62.95	80.3330%	0.4957	0.782	59
	Random Forest	1.43	93.6667%	0.8527	0.935	19
	Random Tree	0.03	88%	0.7216	0.878	34
	REPTree	1.07	81.6667%	0.5844	0.816	55
functions	Voted Perceptron	0.33	59%	0.1782	0.602	123
	SMO	2.45	81.6667%	0.5217	0.793	55
	SGD Text	2.22	66.6667%	0	NA	100
	SGD	4.25	83%	0.5714	0.815	51
	Logistic	286.25	77.6667%	0.4274	0.753	67
	Simple Logistic	15.42	80%	0.4737	0.772	60

As seen in the accuracy column the highest accuracy rate was 94.6667% given by the Random Committee classifier from the group of meta algorithms. On the other hand, the lowest accuracy was given by the Voted Perceptron classifier from the functions group and it was 59%.

The highest and the lowest accuracy rates of the models in each group of algorithms were highlighted with green for the highest rates and red for the lowest. During the implementation of the algorithms, it was observed that several classifiers produced accuracy of 66.6%. This percentage is considered to be a lousy one since all the seizure signals are misclassified, meaning that there is no actual learning performed by the model.

Apart from the accuracy rate of the models, an equally important value was the distribution of errors. Roughly speaking, a good generalization algorithm should present evenly distributed errors within the two classes. However, it should be mentioned that since the healthy instances were more than the seizure ones, a not totally even distribution of errors should be expected. In that way, it is justified the fact that most classifiers presented the most errors in identifying seizure signal.

In the whole testing process, a few algorithms presented equally distributed errors within the classes, accompanied by high accuracy rates. Those algorithms were considered robust to errors and effective on classifying correctly the EEG signal recordings. Such an example was the Random Tree classifier which presented error distribution proportional to the EEG recordings distribution. In this classifier, it was observed that the errors were distributed in 2/3 and 1/3 which is even to the instance distribution since in the dataset there are 200 healthy instances and 100 ictal.

4.4. Classification models

Apart from evaluating the algorithms separately, evaluation of the groups and inquiry for the best classification models was performed. Within this framework, it was observed that a group of algorithms produced better classification results and presented good performance in general. That group was found to be the *trees* which managed accuracy rates between 75.6% and 93.6%. Although the best accuracy (94.6%) belongs to meta algorithms, the trees group is considered the best since it provided better performance in total. That fact was found after averaging the accuracies of each group and it was observed that meta group had an average accuracy of 0.81 and trees group 0.86. Worth mentioning is the fact that the best accuracy result 94.6667% achieved by Random Committee is also based on decision trees since random committee uses the Random Tree as classifier. So, it can be inferred that the tree algorithms perform satisfactorily when it comes to classifying EEG signals. Concerning Random Committee classifier, it should be mentioned that the algorithm produced the longest in-depth tree, namely the size of the tree was 37. All the other decision tree classifiers presented tree size between 9 and 31 apart from the Random Committee which was the deepest.

4.4.1 Tree algorithms description

A general and inclusive definition of decision trees is that they are a graphical representation of specific decisions used when complex branching occurs throughout a decision process. (www.techopedia.com, n.d.) A decision tree is a predictive model based on a branching series of boolean tests that uses specific facts to make more generalized conclusions (www.techopedia.com, n.d.). A decision tree consists of three kinds of nodes which are: *decision*, *chance* and *end* nodes. Every pattern in a decision tree is represented by following the series of paths that were decided starting from the root and reaching the end node within several steps/decisions.

The classifiers provided by weka data mining tool had several characteristics and differences on their approach to classify the EEG recordings. In the following part, a short description of each tree algorithm is provided with the addition of the Random Committee classifier analysis.

Decision stump classifier is a model that consists of a one level decision-tree. (Iba Wayne, 1992) The model contains the root which is directly connected to the leave nodes and makes decisions based on the value of features that are fed to the model one at a time. (Holte Robert, 1993)

Hoeffding tree is an incremental, anytime induction algorithm capable of learning from massive data stream under the condition that the examples fed into the classifier do not change over time.

Hoeffding trees exploit the fact that a small sample can often be enough to choose an optimal splitting attribute. (Geoff Hulten, 2001)

J48 is an extension of the ID3 decision tree algorithm and in weka, it uses the C4.8 algorithm. The J48 takes into account possible missing values, continuous attribute value ranges, decision trees pruning, and derivation of rules. (Gaganjot Kaur, 2014)

LMT classifier uses logistic model trees, which are classification trees that use logistic regression functions at the leave nodes. The algorithm on the specific dataset deals with binary target variables which are nominal. (Niels Landwehr, 2005)

Random Forest classifier creates a group of random trees to be used for the classification. (Breiman, 2001)

Random Tree is a classifier using a tree that takes into consideration k attributes at each node which is chosen randomly. The classifier does not perform pruning and it gives the opportunity to estimate each class probability using a method called back-lifting. (Eibe Frank, -)

REPTree is a fast decision tree learner. The classifier builds a decision tree using information gain/variance and prunes it using reduced-error pruning (with backfitting). (Frank, http://weka.sourceforge.net, -)

Random Committee (from meta algorithms) as the name proposes, is an ensemble of randomizable base classifiers. In this case, the base classifiers were chosen to be random trees. Each of the classifiers contributing on the ensemble is built using the same data but with a random number of seeds. The final predictions of the models are created by averaging the predictions generated by the individual classifiers. (Frank, weka.sourceforge, -)

4.4.2 Experimenting on default options and outcomes

In an effort to improve and further experiment on the results of the classifiers, changes on the default factors were made. In that way, classification models were tested using a different number of cross validation folders (5, 8 and 20- folds) or used training and testing set. In addition, changes on options like the number of iterations made by the classifier, batch of information used in each step, allocated weights etc. In all the above trials it was observed that almost all classifiers presented a slight drop in accuracy or no change at all. Though it should be mentioned that the time taken to build the models fluctuated proportionally to the number of folds of cross-validation or the percentage of data used in the training process.

5. Conclusions

In the above chapters of this thesis a research on algorithms used to classify EEG signals was conducted, along with experiments and trials using Weka 3.8 data mining tool. For the visualizations of the signals and plots, R studio was extensively used, the code of which is displayed in the Appendix. The dataset used in the experiments contained EEG recordings of 300 individuals all adults either healthy or suffering from seizures.

The main purpose of the study was to find models and groups of algorithms that fit EEG data and produce accurate predictions on the class of the data. An initial observation was the fact that the model that emerged from the experiments on weka, was different from the dominant model found in the literature. Specifically, it was found that algorithms like SVM, and ANN which produced very good classification results and were used extensively in classifying EEG signals, belonged to the *functions* group of algorithms. On the other hand, during the experiments conducted using weka, it was observed that the most accurate predictions were produced by *Decision Tree*-relevant algorithms. This deviation could be explained by the difference in the preprocessing part of the data. In the current study, there was no extensive preprocessing performed on the initial dataset and that is a factor that as mentioned above is crucial to the classification results and it is most likely the reason that the group of algorithms producing the best classification results is not the same in the literature and the one found in this thesis. However, there were several examples in which *Tree* algorithms produced very good classification results in the literature. In the study presented above, the classification accuracy achieved by model trials in weka presented notable results especially if we notice that there was minimum preprocess on the actual data and no feature extraction performed in the whole process.

The *Decision Tree* classifiers, presented the best classification accuracy accompanied by short time taken to build each model. However, the fact that the best model was built within a few seconds is a characteristic of the whole group of the *tree* algorithms. The latter is an important observation since it means that a good model that fits the data and presents good predictions, does not need a lot of time to be built, thereby meaning that it is not very complicated. Additionally, the fact that the data were not imposed into normalization or extensive changes gave a major advantage to the whole study. And that is why, the aim of this field of medicine and science is the automation of the processes, and it is important to manage good results and reliable predictions on the initial data coming directly from the EEG machine. *Trees* managed to handle the fact that the signals are almost unique in their whole duration and made that an advantage by classifying correctly the majority of the signals. Finally, it should be mentioned that no overfitting was observed in the results since, apart from the fact that the signals were very different from each other, there were no recurring patterns within the classes that could mislead the *tree* algorithms.

Further studies and extensive experiments on tree models, aiming at improvements on the existing algorithms or the creation of new ones, along with further preprocessing techniques would be future research goals. Also, in future studies, it would be interesting to find dominant features and discard others that may cause overfitting in the classification process.

Bibliography

References

Ahnaf Rashik Hassan, A. S. (2016). Automatic identification of epileptic seizures from EEG signals using linear programming boosting. *Computer methods and programs in biomedicine.*, 65-77.

Alexandros T. Tzallas, M. G. (2009). Epileptic Seizure Detection in EEGs Using Time–Frequency Analysis. *IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE*, 703-710.

Andrew B. Gardner, A. M. (2006). One-Class Novelty Detection for Seizure Analysis from Intracranial EEG. *Journal of Machine Learning Research*, 1025 – 1044.

Breiman, L. (2001). Random Forests. *Machine Learning.* , 5-32.

Daubechies, I. (1992). Ten lectures on wavelets. Philadelphia: Society for Industrial and Applied Mathematics.

Eibe Frank, R. K. (-). *weka.sourceforge.net*. Retrieved from weka sourceforge: http://weka.sourceforge.net/doc.dev/weka/classifiers/trees/RandomTree

Emina Alickovic, J. K. (2018). Performance evaluation of empirical mode decomposition, discrete wavelet transforms, and wavelet packed decomposition for automated epileptic seizure detection and prediction. *Biomedical Signal Processing and Control*, 94-102.

Enamul Kabir, S. Y. (2016). Epileptic seizure detection from EEG signals using logistic model trees. *Brain Informatics*, 93-100.

Frank, E. (-). *http://weka.sourceforge.net*. Retrieved from weka.sourceforge.net: http://weka.sourceforge.net/doc.dev/weka/classifiers/trees/REPTree.html

Frank, E. (n.d.). *http://weka.sourceforge.net*. Retrieved from http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/SGD.html: http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/SGD.html

Gaganjot Kaur, A. C. (2014). Improved J48 Classification Algorithm for the Prediction of Diabetes. *International Journal of Computer Applications*.

Geoff Hulten, L. S. (2001). Mining time-changing data streams. *ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining*, (pp. 97-106).

Gotman, J. (1990). The use of computers in analysis and display of EEG and. *New York: Raven*, 51-83.

Hojjat Adeli, Z. Z. (2003). Analysis of EEG records in an epileptic patient using wavelet transform. *ELSEVIER*, 19.

Holte Robert. (1993). *Very Simple Classification Rules Perform Well on Most Commonly Used Datasets.*

Iba Wayne, L. P. (1992). Induction of One-Level Decision Trees. *Proceedings of the Ninth International Conference on Machine Learning*, pp. 233–240.

Javad Birjandtalab, M. B. (2017). Automated seizure detection using limited-channel EEG and non-linear dimension reduction. *Science Direct*, 49-58.

Lina Wang, W. X. (2017). Automatic Epileptic Seizure Detection in EEG Signals Using Multi-Domain Feature Extraction and Nonlinear Analysis. *MDPI*, 17.

Ling Guo, D. R. (2010). Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. *Journal of Neuroscience Methods* , 101-109.

Lorena Orosco, A. G. (2016). Patient non-specific algorithm for seizures detection in scalp EEG. *Computers in Biology and Medicine*, 128-134.

Lüders HO, E. J. (1993). The epileptogenic zone: general principles. *Surgical Treatment of the Epilepsies*, 53.

Manish N. Tibdewala, H. R. (2017). Multiple entropies performance measure for detection and localization of multi-channel epileptic EEG. *Biomedical Signal Processing and Control*, 158-167.

Mayo Clinic . (n.d.). Retrieved from https://www.mayoclinic.org/: https://www.mayoclinic.org/diseases-conditions/seizure/symptoms-causes/syc-20365711

Md. Kamrul Hasan, M. A. (2017). Prediction of Epileptic Seizure by Analyzing Time Series EEG Signal Using k-NN Classifier. *Applied Bionics and Biomechanics*.

Md. Mamun or Rashid, M. A. (2017). Epileptic Seizure Classification using Statistical Features of EEG Signal. 6.

Miller, K. (2003). *https://www.thefreedictionary.com/*. Retrieved from medical-dictionary.thefreedictionary.com: https://medical-dictionary.thefreedictionary.com/temporal+lobe

Mingyang Li, W. C. (2017). Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble. *Biomedical signal processing and control*, 357-365.

Mohammed Diykh, Y. L. (2017). Classify epileptic EEG signals using weighted complex networks-based community structure detection. *Expert Systems With Applications*, 87-100.

Mosby. (2009). *medical-dictionary.thefreedictionary.com*. Retrieved from Hippocampal Formation: https://medical-dictionary.thefreedictionary.com/hippocampal+formation

Mutlu, A. Y. (2018). Detection of epileptic dysfunctions in EEG signals using Hilbert vibration decomposition . *Biomedical Signal Processing and Control*, 33–40.

Niels Landwehr, M. H. (2005). Logistic Model Trees. *Machine Learning*, 161-205.

Nurettin Acır, ˙. Ö. (2005). Automatic Detection of Epileptiform Events in EEG by a Three-Stage Procedure Based on Artificial Neural Networks. *IEE*.

NYU Langone Medical Center. (n.d.). Retrieved from https://med.nyu.edu/: http://www.med.nyu.edu/cec/epilepsy/types/primary_generalized.html

Ocak, H. (2008). Optimal classification of epileptic seizures in EEG using wavelet analysis and genetic algorithm. *Signal Processing*, 1858–1867.

Ocak, H. (2009). Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Systems with applications*, 2027-2036.

Ozan Kocadagli, R. L. (2017). Classification of EEG signals for epileptic seizures using hybrid artificial neural networks based wavelet transforms and fuzzy relations. *Expert systems with applications*, 419-434.

Polat, K. ,. (2007). Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. *blaaaaaa*, 1017-1026.

Rajeev Sharma, R. B. (2015). Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions. *Expert Systems with applications 42*, 1106-1117.

Rajendra Acharya. U, F. H. (2015). Application of entropies for automated diagnosis of epilepsy using EEG signals: A review. *ScienceDirect*, 85–96.

Rajendra Acharya.U, F. M.-H. (2012). Automated diagnosis of epileptic EEG using entropies. *Biomedical Signal Processing and Control*, 401-408.

Ralph G. Andrzejak, K. L. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *The American Physical Society*, 1-8 DOI: 10.1103/PhysRevE.64.061907 .

Ralph Meier, H. D.-B. (2008). Detecting Epileptic Seizures in Long-term Human EEG: A New Approach to Automatic Online and Real-Time Detection and Classification of Polymorphic Seizure Patterns. *Journal of Clinical Neurophysiology*.

Ramy Hussein, M. E. (2018). Robust detection of epileptic seizures based on L1-penalized robust regression of EEG signals. *Expert Systems With Applications*, 153-167.

Sample Entropy. (n.d.). Retrieved from https://physionet.org/physiotools/sampen/

Shivnarayan Patidar, T. P. (2017). Detection of epileptic seizure using Kraskov entropy applied on tunable-Q wavelet transform of EEG signals. *Biomedical Signal Processing and Control*, 74-80.

Subasi, A. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications*, 1084-1093.

Suryannarayana Chandaka, A. C. (2009). Cross-correlation aided support vector machine classifier for classification of EEG signals. *Expert Systems with Applications* , 1329–1336.

Tao Zhang, W. C. (2018). Generalized Stockwell transform and SVD-based epileptic seizure detection in EEG using random forest. *Biocybernetics and Biomedical Engeneering* , 519-534.

Tapan Gandhi, B. K. (2010). Expert model for detection of epileptic activity in EEG signature. *Expert Systems with Applications*, 3513–3520.

Tedrus GM, F. L. (2012). Epilepsy with onset at over 50 years of age: clinical and electroencephalographic haracteristics. *Arq Neuropsyquiatrics*.

U. Rajendra Acharya, S. L. (2017). Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in Biology and Medicine*, 1-9.

www.techopedia.com. (n.d.). Retrieved from techopedia: https://www.techopedia.com/definition/28634/decision-tree

Zakareya Lasefr, S. S. (2017). Epilepsy Seizure Detection Using EEG signals. 8.

Zamir, Z. R. (2016). Detection of epileptic seizure in EEG signals using linear least squares preprocessing. *Computer methods and programs in biomedicine*, 95-109.

Zeynab Mohammadpoory, M. N. (2017). Epileptic seizure detection in EEGs signals based on the weighted visibility graph entropy. *Seizure*, 202-208.

Zhicheng Cui, W. C. (2016). Multi-Scale Convolutional Neural Networks for Time Series Classification. p. 10.

Appendix

The R code used for the visualizations in chapter 3 and for further exploration of the data is distributed below. It should be mentioned that the datasets used were not in the same structure as the ones loaded into the Weka 3.8 data mining tool.

```
# remove loaded values

rm(list = ls())

install.packages("stringr");

install.packages("installr");

if(!require(installr)) { install.packages("installr"); require(installr)} #load / install+load installr

library(installr)

# step by step functions:

check.for.updates.R() # tells you if there is a new version of R or not.

install.R() # download and run the latest R installer

copy.packages.between.libraries() # copy your packages to the newest R installation from the one version before it (if ask=T, it will ask you between which two versions to perform the copying)

trace(utils:::unpackPkgZip, edit=TRUE)

install.packages("xlsx");

install.packages("rJava");

install.packages("ggplot2");

install.packages("plyr");

install.packages("dplyr");

install.packages("tidyverse");

install.packages("DataExplorer");

install.packages("dplyr");

install.packages("corrplot");

library(dplyr)

library(xlsx)

library(ggplot2)

library(plyr)
```

```
library(dplyr)
library(stringr)
library(DataExplorer)
library(tidyverse)
library(corrplot)
fdata<-data.frame(F_DATASET_NON_SEIZURE)
odata<-data.frame(O_DATASET_NON_SEIZURE_WITH_CLOSED_EYES)
sdata<-data.frame(S_DATASET_SEIZURE)
#shows the min max values of the columns
summary(fdata)
summary(odata)
summary(sdata)
#shows which vlues are more common and the distribution of the data
hist(fdata$F060)
hist(odata$O001)
hist(sdata$S001)
#line plot of F the 3 files that contain the mix,max and a random choice F001
ggplot(data=fdata, aes(x=ID, y=F001, group=1)) +
  geom_line()+
  geom_point() + ylim(-2000, 2100)+ xlim(0,4097)
ggplot(data=fdata, aes(x=ID, y=F009, group=1)) +
  geom_line()+
  geom_point() + ylim(-2000, 2100)+ xlim(0,4097)
ggplot(data=fdata, aes(x=ID, y=F060, group=1)) +
  geom_line()+
  geom_point() + ylim(-2000, 2100)+ xlim(0,4097)
#matplot(fdata$F001)
#line plot of the O file containing min,max and rondom choise O001
ggplot(data=odata, aes(x=ID, y=O001, group=1)) +
  geom_line()+
```

```
  geom_point() + ylim(-2000, 2100)+ xlim(0,4097)
ggplot(data=odata, aes(x=ID, y=O017, group=1)) +
  geom_line()+
  geom_point() + ylim(-2000, 2100)+ xlim(0,4097)


ggplot(data=odata, aes(x=ID, y=O100, group=1)) +
  geom_line()+
  geom_point()+
  ylim(-2000, 2100)+ xlim(0,4097)
#matplot(odata$O001)
#matplot(odata$O017)
#matplot(odata$O100)
#line plot of the S file containing min,max and rondom choise S001
ggplot(data=sdata, aes(x=ID, y=S001, group=1)) +
  geom_line()+
  geom_point() + ylim(-2000, 2100)+ xlim(0,4097)
ggplot(data=sdata, aes(x=ID, y=S010, group=1)) +
  geom_line()+
  geom_point() + ylim(-2000, 2100)+ xlim(0,4097)
ggplot(data=sdata, aes(x=ID, y=S011, group=1)) +
  geom_line()+
  geom_point() + ylim(-2000, 2100)+ xlim(0,4097)
#shows the concentration of frequrences detected in the different files
f <- density(fdata$F001)
plot(f)
o<-density(odata$O001)
plot(o)
s<-density(sdata$S001)
plot(s)
#correlation matrixw within folders
```

```
l<-cor(fdata)
corrplot(l, method="color", type = "upper")
k<-cor(odata)
corrplot(k, method="color", type = "upper")
m<-cor(sdata)
corrplot(m, method="color", type = "upper")
#correlation between different random folders columns
cordata<-data.frame(Book1)
n<-cor(cordata)
corrplot(n, method="color", type = "upper")
#correlation within multiple columns from all the files
mixed<-data.frame(mixed_data)
o<-cor(mixed)
corrplot(o,method="color", type = "upper")
```